# Network Slicing in 5G – A Paradigm Shift

Dr. Debabrata Das
Professor and HP Chair
IIIT-Bangalore

# Agenda

- Introduction
- 5G Network  Architecture
- Concept of Network Slicing
- Realization in Core Network, RAN and Transport
- Challenges in 5G Network Security
- 
- Network Slicing Studies @ IIIT-Bangalore
- Summary

# Industry 4.0 Demands

World has becomes ever more digitally and globally connected with every industries embracing the ICT innovations

INDUSTRY 4.0

**Manufacturing:**
Smart Factory with IoT and Automa`ıtion, Robots

**Healthcare:**
Smart Health gadgets
Patient monitoring
Cloud based Data Storage

**Media & Entertainment:**
Immersive forms
Digital media transmission and reception

**Financial:**
E-wallets, surge of online payment, mobile based transactions

**Public Safety:**
CCTV Surveillance, Cyber-Security

**Automotive:**
Autonomous driving, Connected cars, office-on-move

**Transport:**
Emission and System monitoring, Tracking, Internet accessibility

**Energy:**
Smart meters, Broadband over Power Lines, System monitoring

**Education:**
Digital classrooms, online courses

# Present 4G Challenges?

- 4G standard started in 2000 → at that time performance requirements for massive IoT, Driver less car, robotic operation, etc. were not visualized → 4G's multiple releases tried to adopt
- Hence, current mobile network cannot give a right user experience of Industry 4.0 uses due to lack of throughput, latency, reliability, availability
- New Wireless Technology like 5G is not just about making the throughput larger than LTE/LTE-A
- All about offering user experiences and inclusion of all Industry 4.0 sectors which haven't been included in LTE/LTE-A (like, ultra low latency, massive IoT,…)

# Main requirement of 5G

Hence, 5G is expected to provide a great variety of services in broadly following 3 categories; e.g.,

- **Enhanced mobile broadband (eMBB)**
- **Massive machine type communication (mMTC),**
- **Ultra reliability low latency communication (URLLC)**

# Services and Requirements Characteristics [2]

| Characteristics | mMTC | URLLC | eMBB |
| --- | --- | --- | --- |
| **Availability** | Regular | Very High | Regular (baseline) |
| **E2E latency** | Not highly sensitive | Extremely sensitive | Not highly sensitive |
| **Throughput type** | Low | Low/med/high | Medium |
| **Frequency of Xfers** | Low | High | High |
| **Density** | High | Medium | High |
| **Network coverage** | Full | Localized | Full |

# Examples:

## India centric : Aadhar based Authentication in Airport

- Aadhar Centric Matching→ End-to-end Delay of a search < 1 Sec (1000ms)
  - The above has: Network delay + Search Delay in Data Centre (DC) + Display
  - Network delay < 300ms
- Network Delay = 2 X (user device first mile delay + *N* number hops to DC + Last mile network in data centre) < 300 ms

## Global Requirements: Driver Less car, or sub-millisecond tactile services (remote surgery)

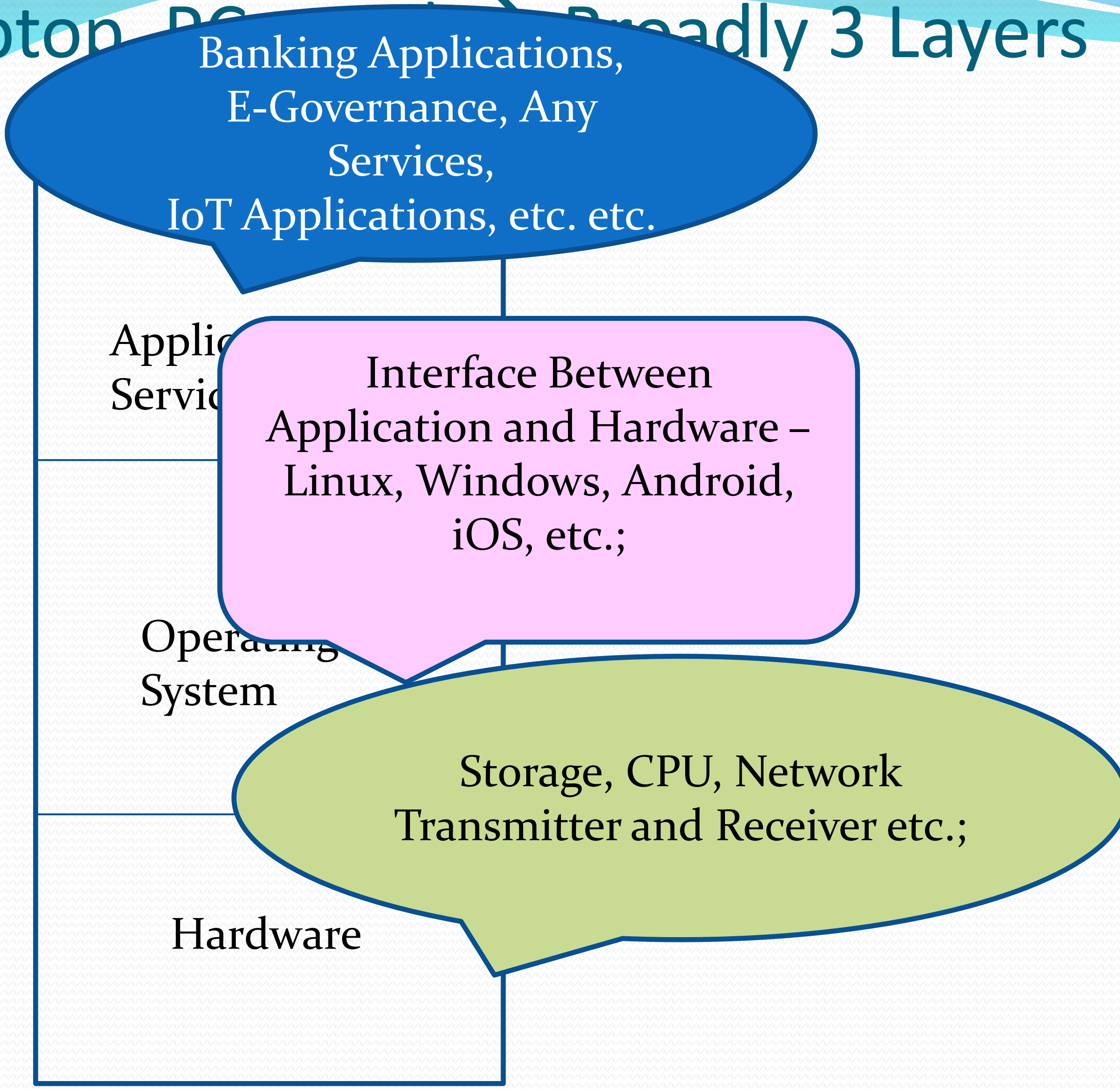
- Driver Less Car: camera to break = 1 ms
- Remote Surgery < a FEW ms (assume 10 ms)

Hence, Service centric Architecture:
Application, Transport, Network Layer, Data-link and PHY will be defined and optimized end-to-end → Multiple TCP/UDP, Routing, MAC and PHY innovation and OS, hardware optimization in a system

# Where is the challenge for the previous slide requirements?

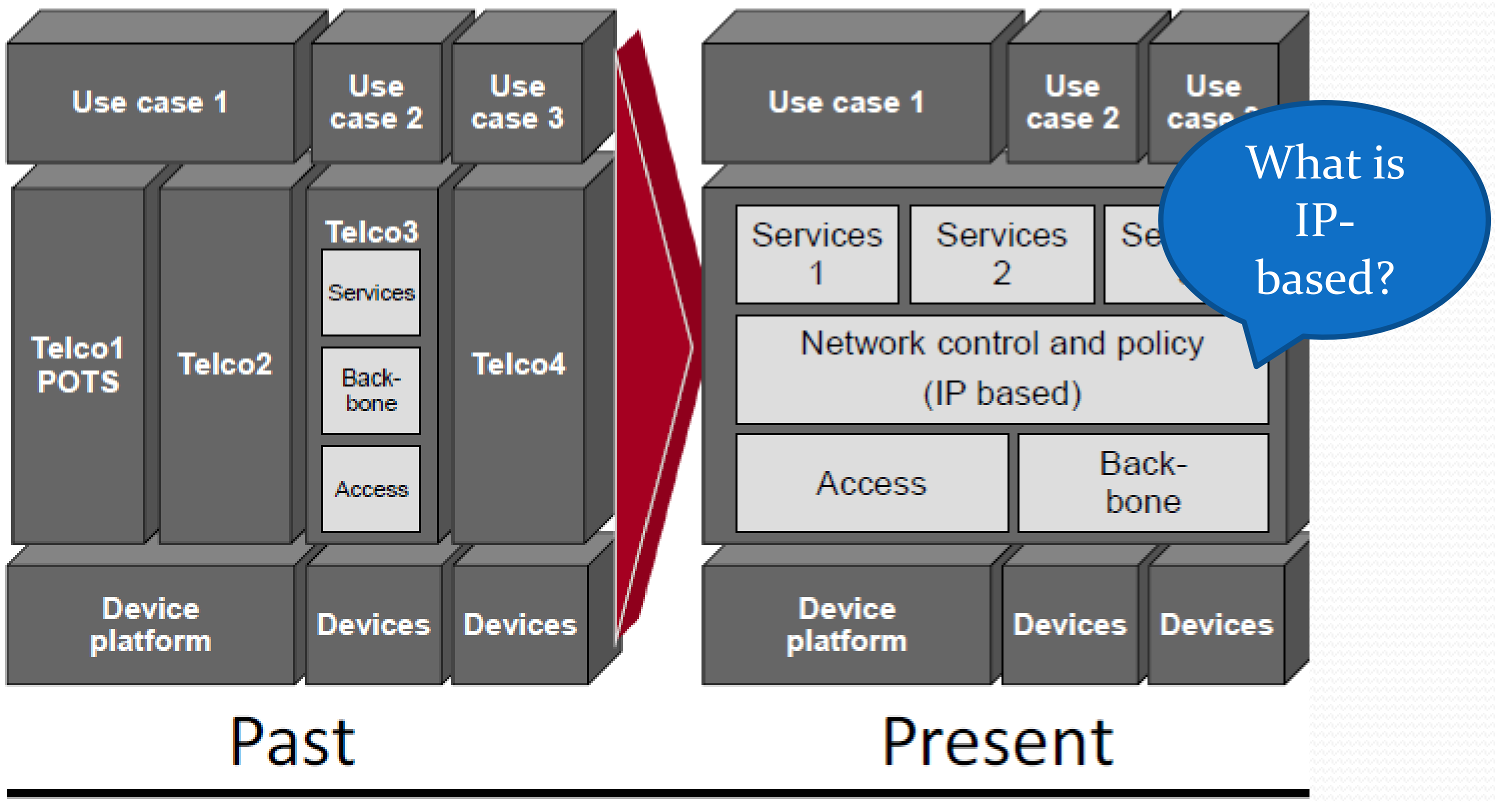
Let us start from fundamentals of IT/Network Systems!

# Any IT System (Router, Switch, Storage, IoT, Mobile, Laptop, PC, ...) – Broadly 3 Layers



# Evolution of the network infrastructure

From Silos Architecture to Monoliths  Architecture



# Monolithic Architecture (Present)

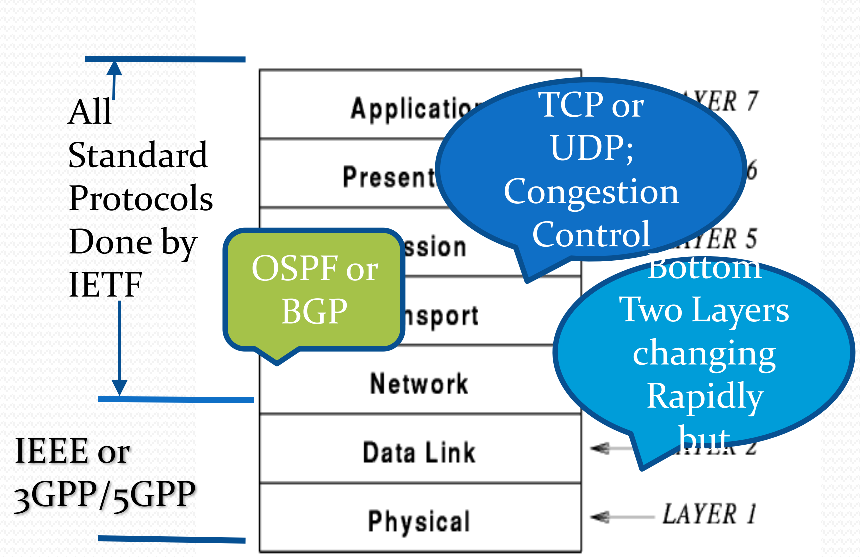
- Discarded the proprietary based protocol systems and hardware of Silo
- Since it was IP based, interworking of the applications became simple → any user of one network can call the other user of another network
- **Thus removing the gateway type of products**
- **So the concept of services came into existence** (voice, video, email etc.,)
- Though the telecom operator have same products, at the top down it looks like a single monolithic product block.
- **Main drawback of this approach is that each telecom operator's configuration had an impact on the packet exchange across the networks**
  - ❖ **Guaranteeing the packet exchange requires complicated algorithms with cross layer knowledge and implementations**
  - ❖ Low latency services CANNOT be satisfied by one-size-fits-all monolithic architecture

# Layering Structure of Network

- Why Layering Structure?

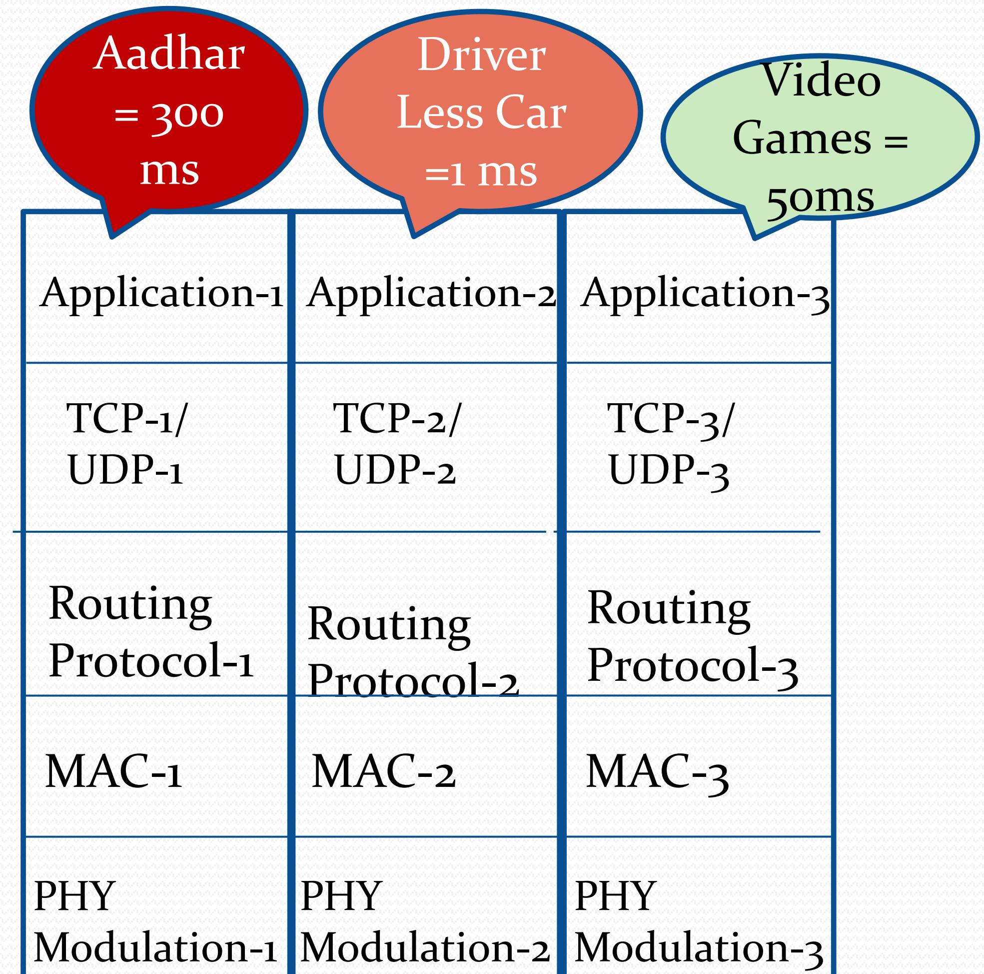
→ Model suggests dividing the network protocol into *layers*, each of which solves part of the network communication problem

- ISO Standard Network has 7 Layers

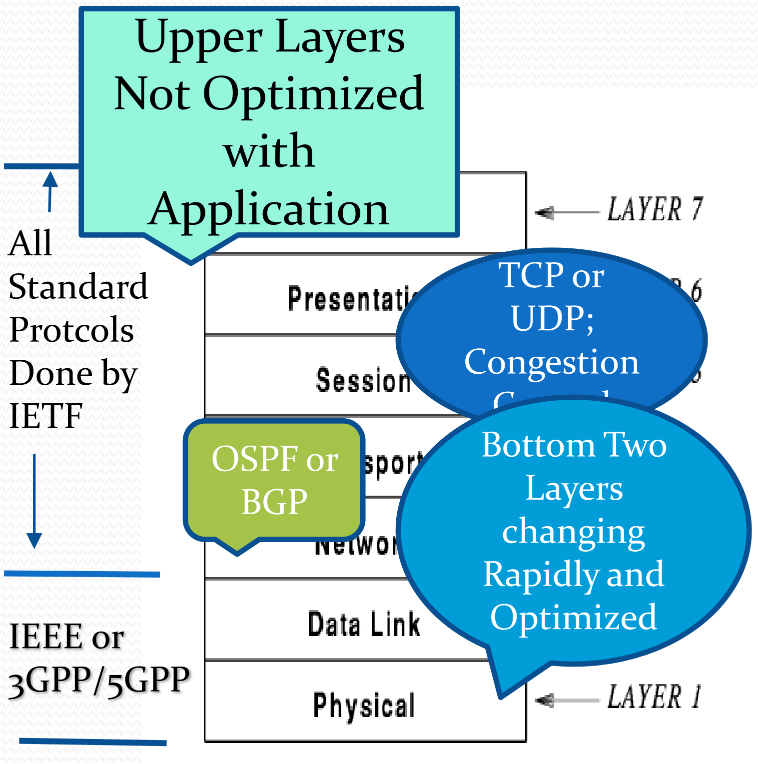


# Layering Structure of Network: Vertical Optimization

Future: Source UE/IoT and
Final Destination UE/IoT or Server

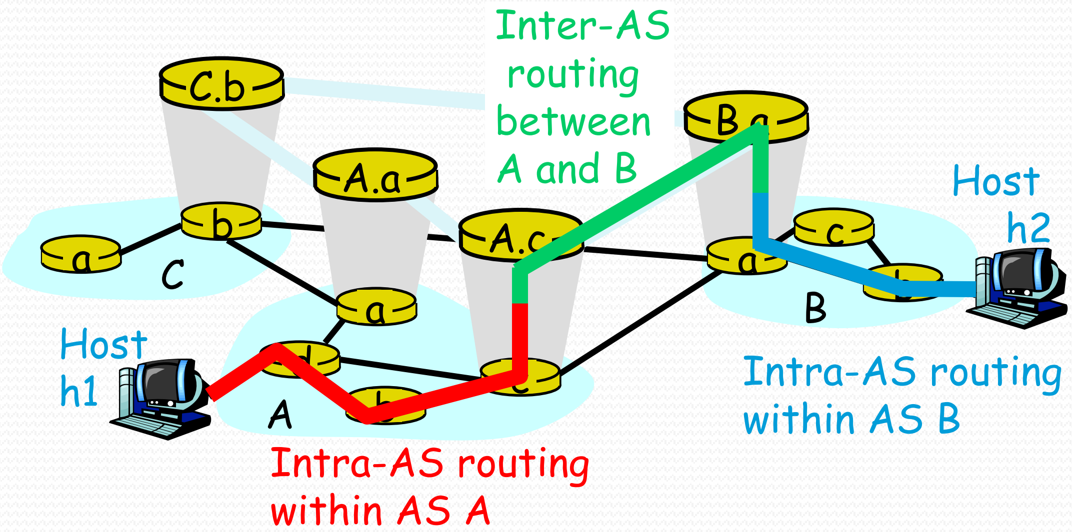



What about Intermediate nodes Between
Source and Final Destination?




Present: Source UE/IoT and
Final Destination UE/IoT or Server

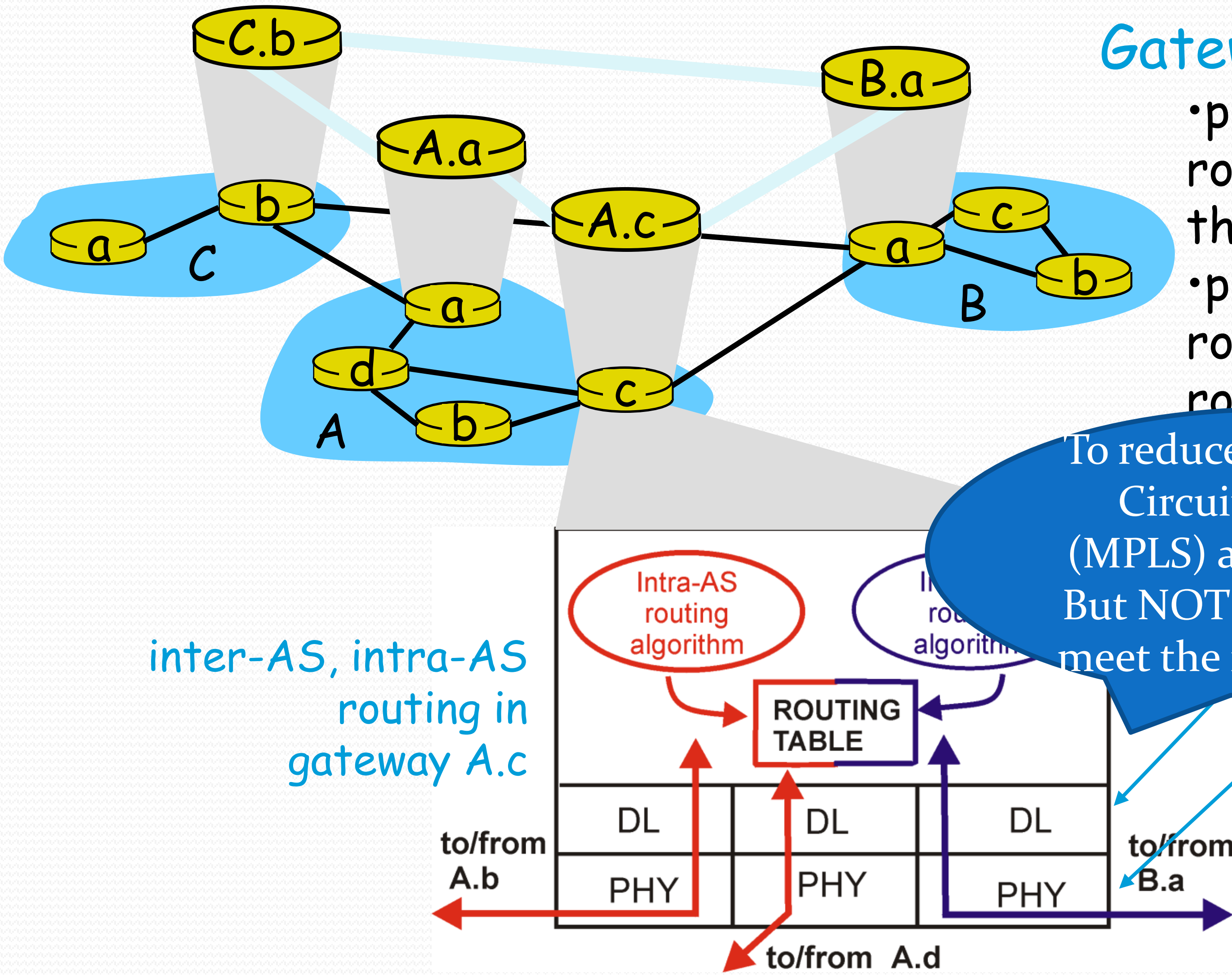
# Intra-AS and Inter-AS routing: Horizontal View




Delay Calculations in each Hop: (Queuing + Processing + Transmission + Propagation)

1. UE in Bangalore and Server in California (**Only Propagation Delay**):
5 micro Sec/1 km X 14000 kms = 70 milli Sec. (one direction)
2. UE in Bangalore and Aadhar Server in Delhi (Only propagation Delay):
5 micro Sec/1 km X 2000 kms = 10 ms (one direction)

# Intra-AS and Inter-AS routing in Gateway Router

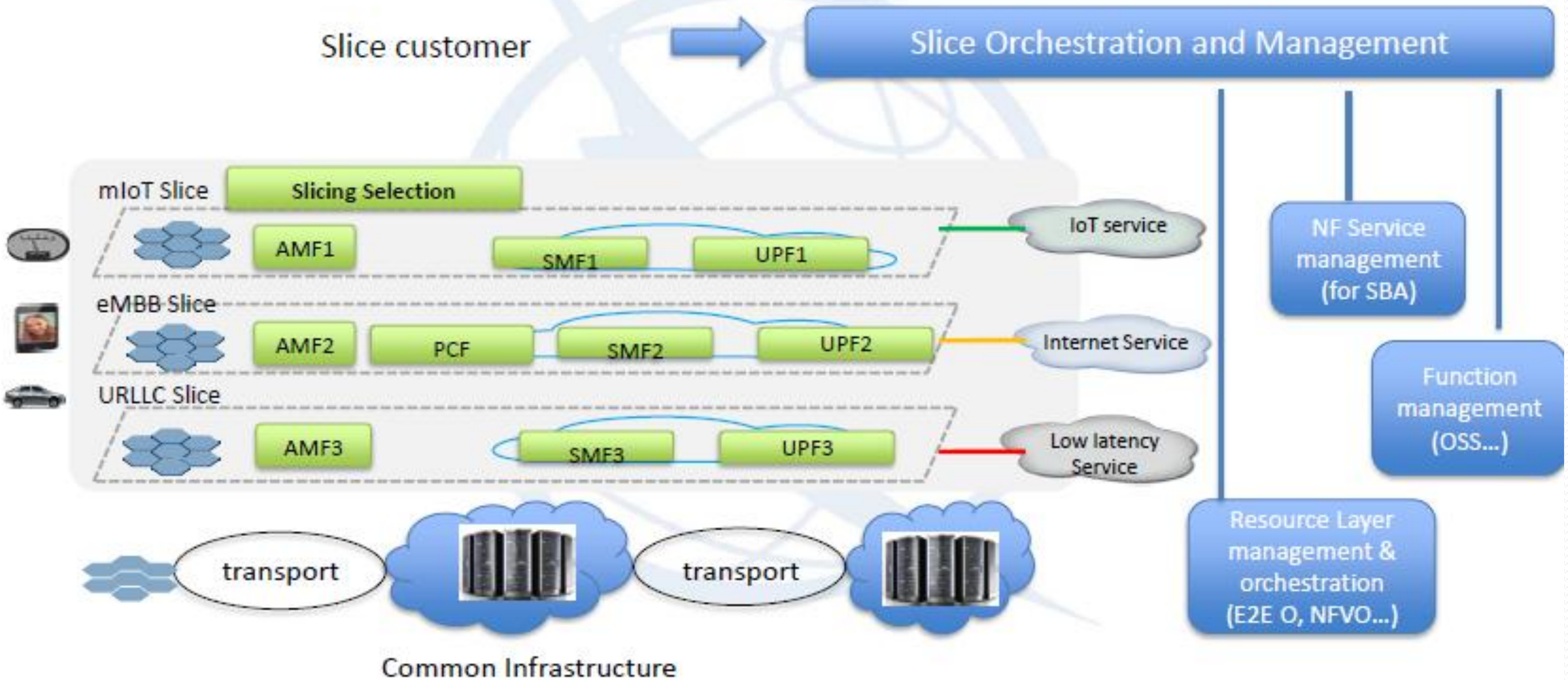



Gateways:

- perform inter-AS routing amongst themselves
- perform intra-AS routers with other routers in their

To reduce delay Virtual Circuit switching (MPLS) and now SDN: But NOT enough yet to meet the required delay

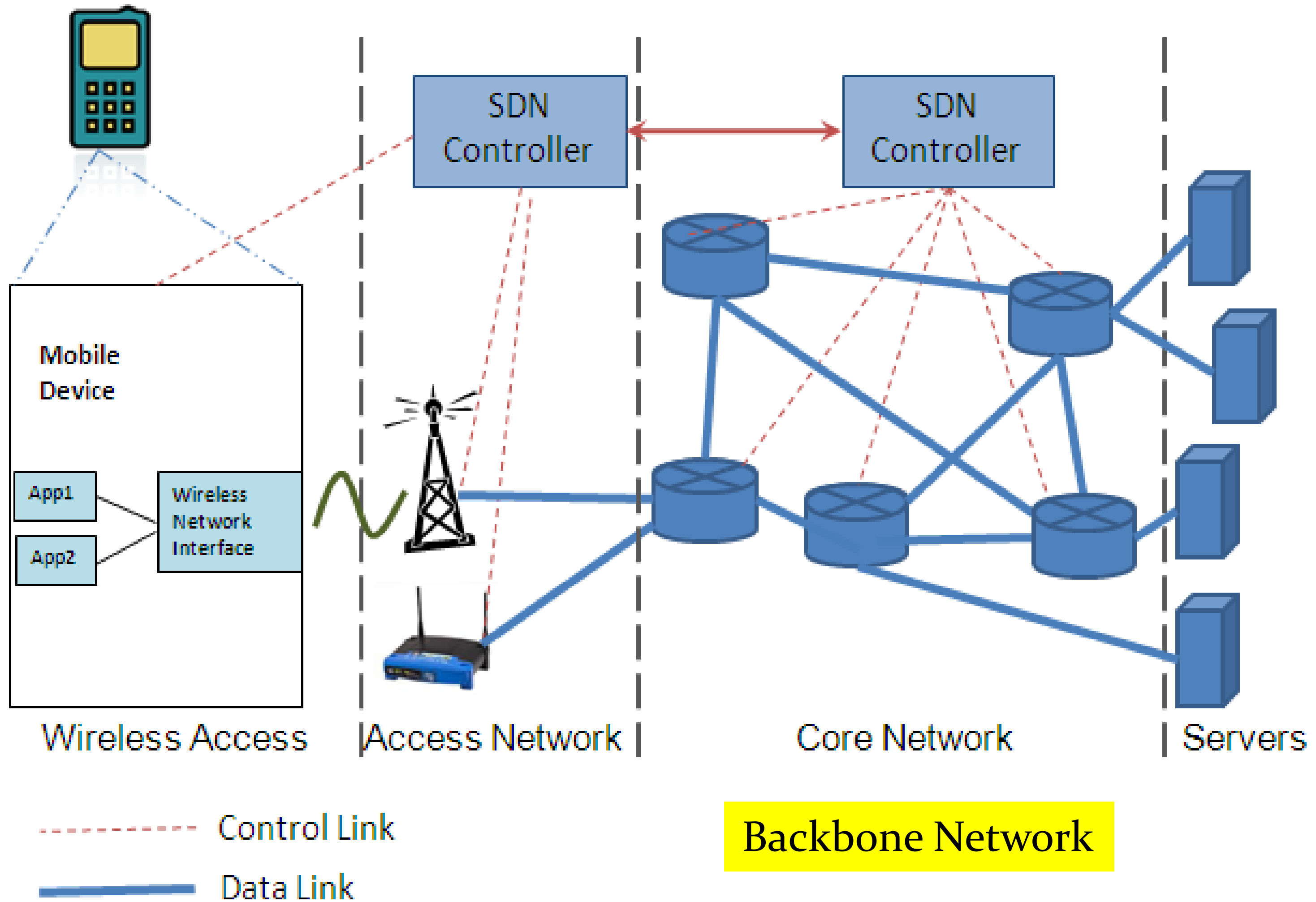
# Network Slicing of 5G Architecture



# One-size-fits-all architecture will change to Service Oriented Slicing

- Vertical:
  - Any service has the same set of protocols (TCP, UDP, Congestion Control, Routing etc.), OR, a limited set of protocols (4G, 4G-Advance, WiFi) and hardware resources
-
- Horizontal:
  - Limited Inter-domain QoS mapping for an end-to-end communication

# Example: SDN/NFV will help for slicing and optimization




Dibakar Das, J. Bapat and Debabrata Das, "A Dynamic QoS Negotiation Mechanism Between Wired and Wireless SDN Domains", *IEEE Transaction on Network and Service Management*, Volume: 14, Issue: 4, December 2017.

# Direction for 5G

Architecture = Topology + Protocols/Communication Rules

# Direction of 5G Systems and Architecture

- World wide various standard bodies working on 5G Systems requirements and exact performance goal
- However, industry and academia agreeing TWO following complementary views, which is driving the present research and industry on 5G
  - **Evolutionary View**
  - **Service Oriented View**

# Evolutionary View

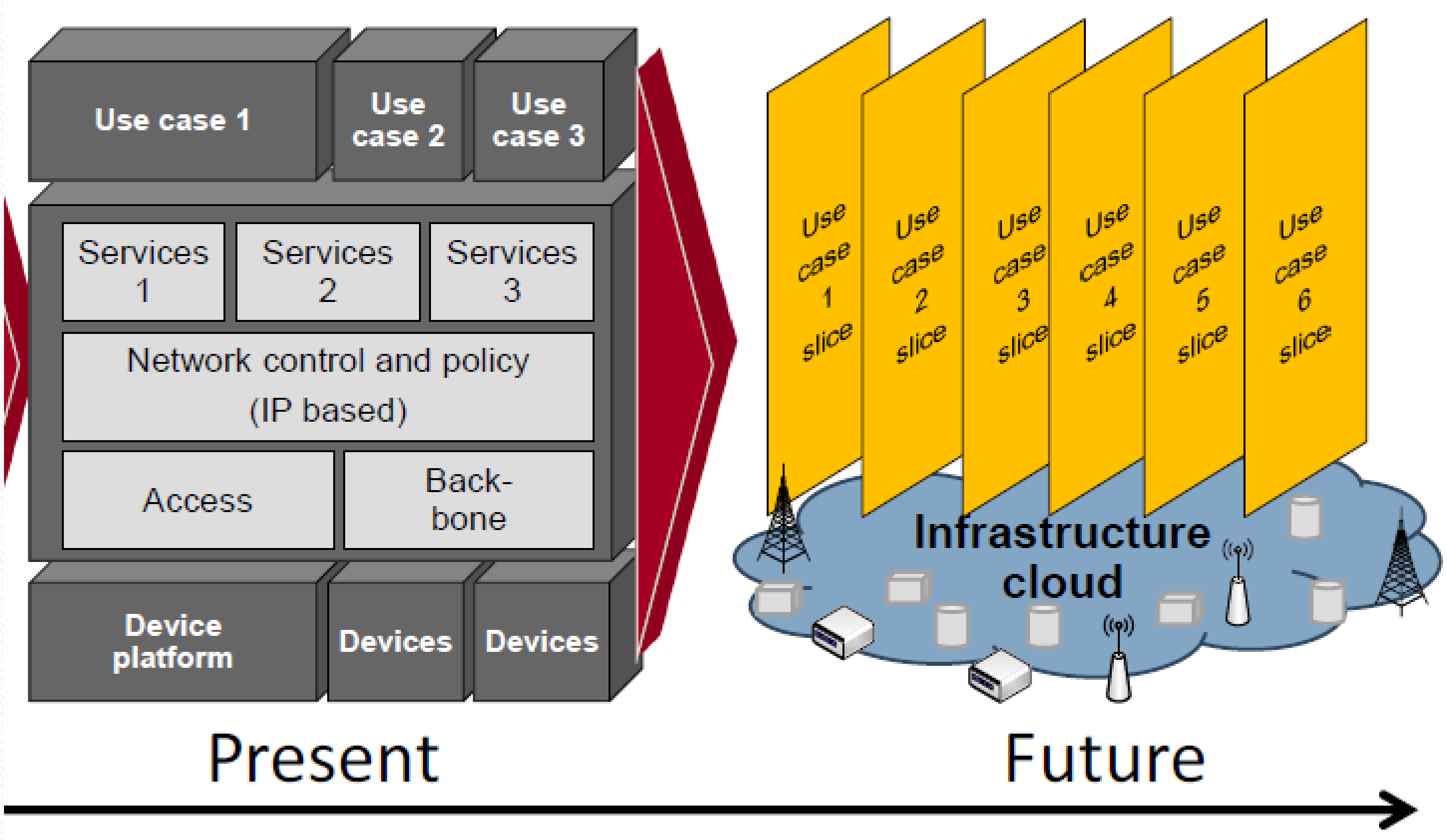
- Natural Transition from 4G-Advanced (4.5G) to 5G
- Focusing on significantly improving the efficiency of mobile networks
  - 1000X traffic volume
  - 100X devices (UE+ IoT + other devices on Network over next 10 years)
  - 100X Throughput
  - 100X delay reduced
  - Spectrum efficiency (2G to 4G increased by 30 times)

# Service-Oriented View

- Envisioned to cater wide range of services → from traditional human centric to Machine-type communication (as discussed in previous example, driverless car, robotic operation etc.)
- It requires same physical infrastructure of network to take different forms depending on the service requirements → naturally to the notion of network slicing
- Network Slicing → Service-oriented view required for radical rethinking of existing mobile and backbone network architecture with
  - More flexible and programmable network systems
  - As mentioned before → above possible by leveraging new technologies like Software Defined Network (SDN) and Network Function Virtualization (NFV), Cloud and Edge Computing

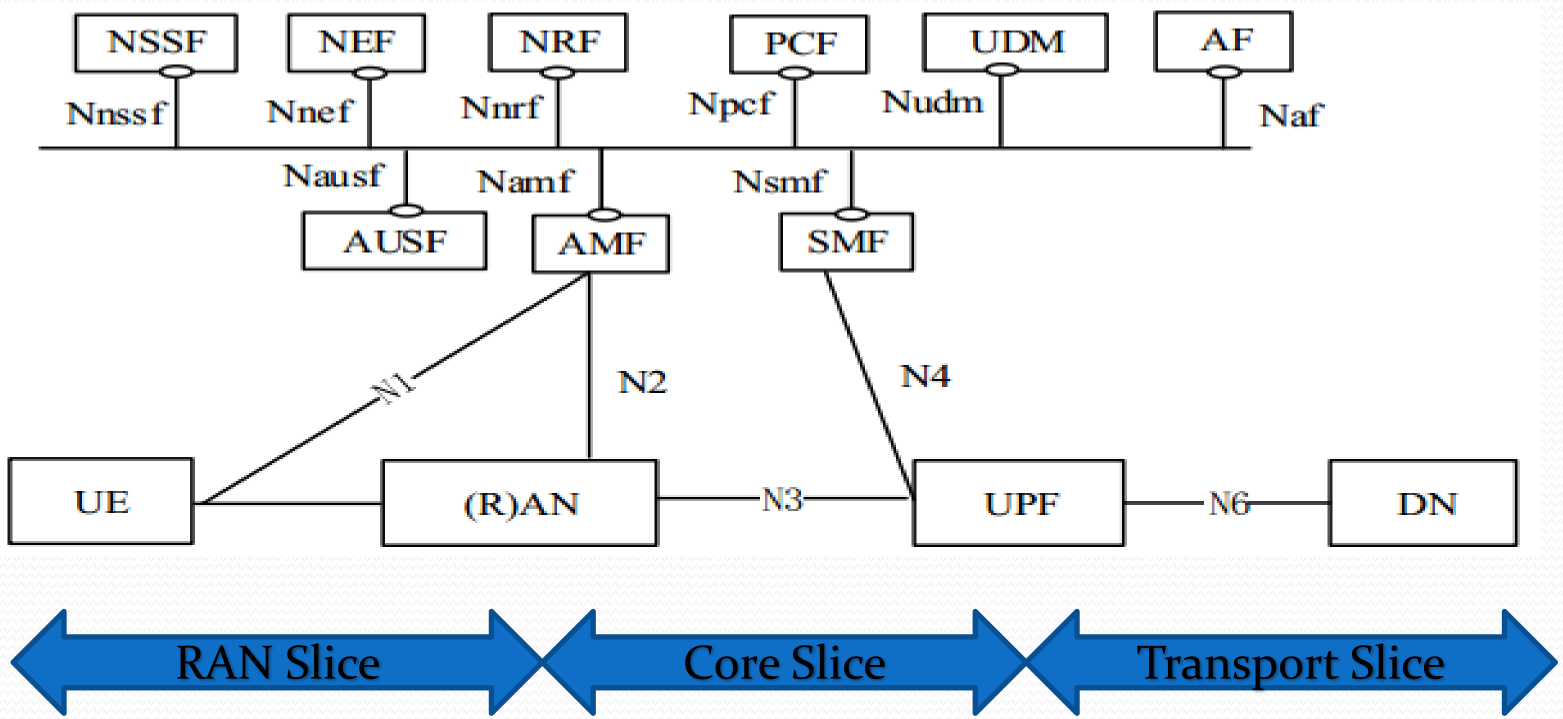
# Realizing Network Slicing (contd.,)

From Monoliths  Architecture to Network Slicing



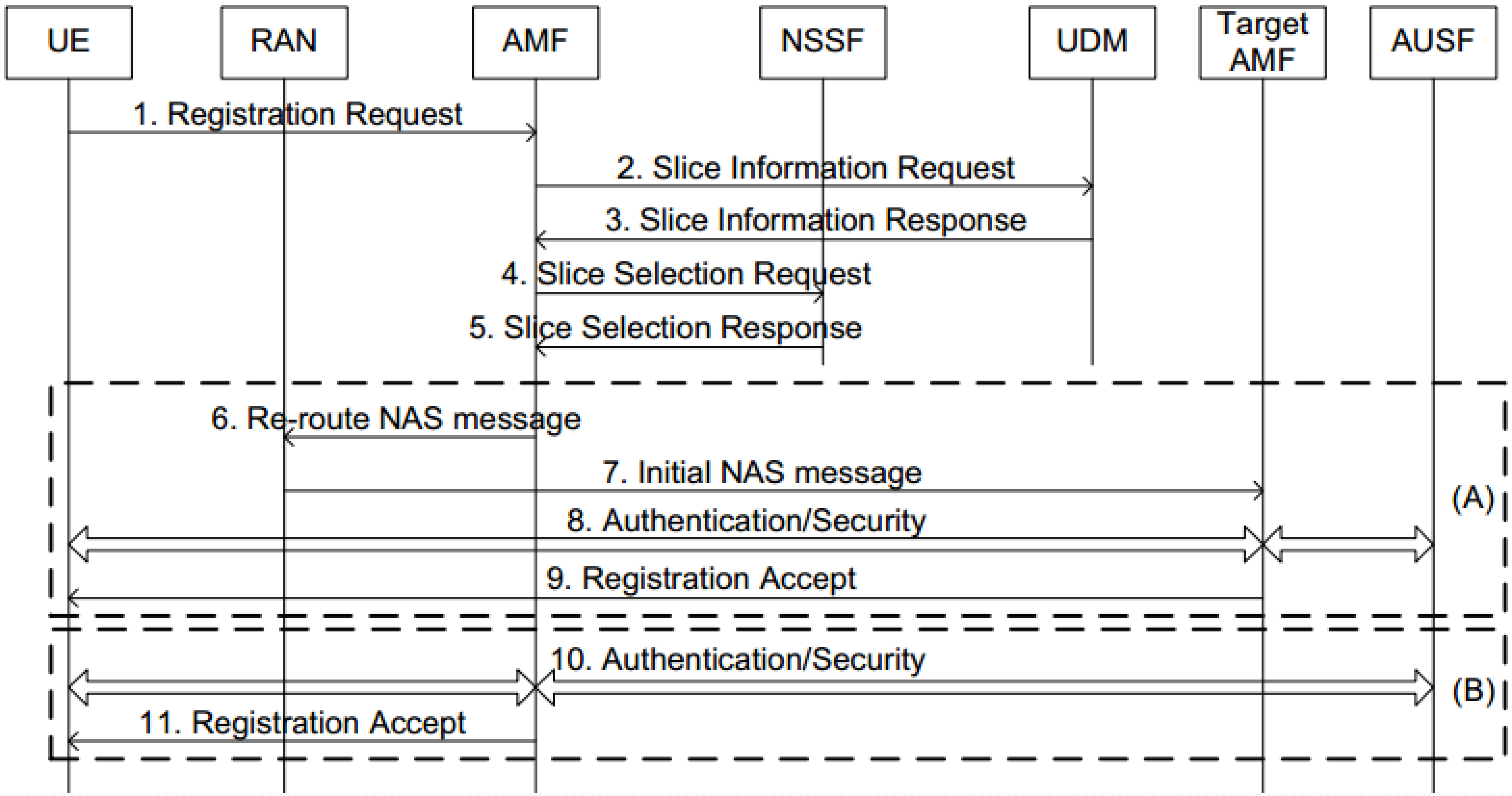

ERLEBEN, WAS VERBINDET.

# RAN—Core—Transport Slicing [2]



- Authentication Server Function (AUSF): supports authentication server function.
• Access and Mobility Management Function (AMF): access control, mobility control, transparent proxy for routing
SM message.
• Unified Data Repository (UDR): storage and retrieval of data by the UDM, PCF or NEF.
• Unstructured Data Storage Network Function (UDSF): storage and retrieval of information as unstructured data by
any NF.
• Network Exposure Function (NEF): expose the services and capabilities provided by 3GPP network functions.
• NF Repository Function (NRF): maintains NF profile, supports service discovery.
• Policy Control function (PCF): decides the policy and provides them to the control plane function.
• SMF (Session Management Function): manages the PDU session e.g. PDU session establishment, modify and
release.
• Unified Data Management (UDM): authentication credential processing, access authorization, registration/ mobility management and subscription management.
• User plane Function (UPF): handles the user plane traffic, e.g. traffic routing & forwarding, traffic inspection and usage reporting, handling.
• Application Function (AF): interacts with the 3GPP Core Network (CN) to provide services.
• Network Slice Selection Function (NSSF): selects the NSI, determines the allowed network slice selection assistance information (NSSAI) and AMF set to serve the UE.

# Registration Procedure in Core N/W



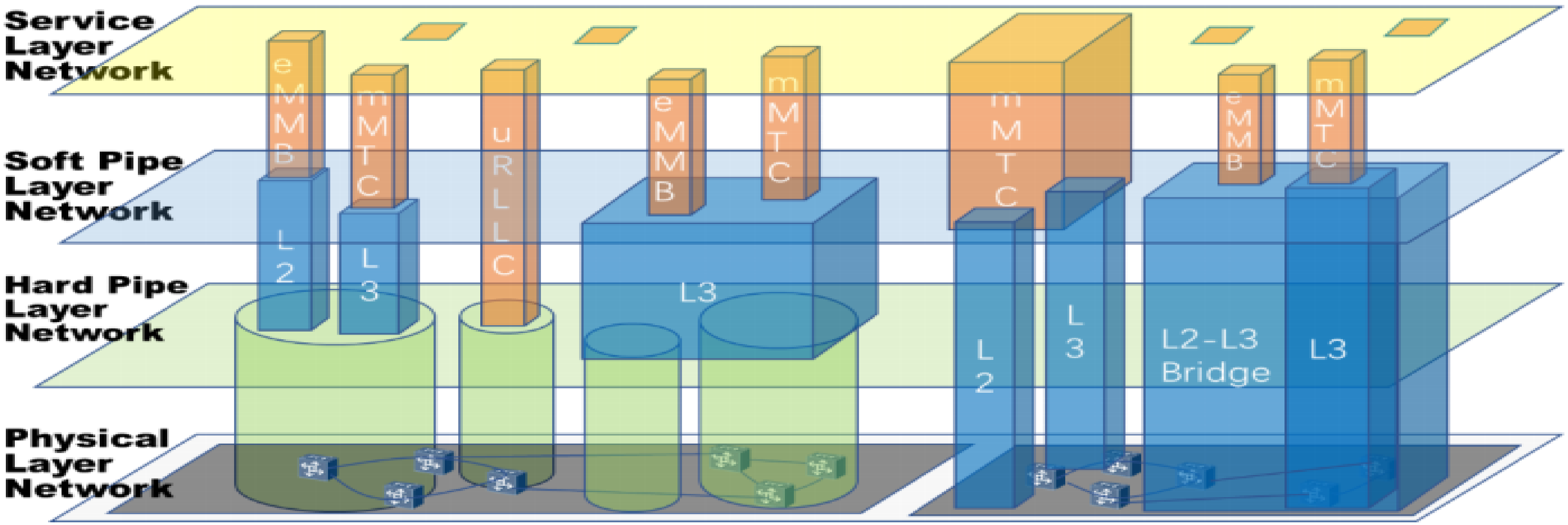
# Steps of Registration in Previous Slides

- 1. UE sends Registration Request to AMF. The message includes requested NSSAI.
  2. AMF sget subscription data from UDM.
  3. UDM responds subscription data to AMF. The subscribed NSSAI includes subscribed S-NSSAI.
  4. AMF sends Slice Selection Request to NSSF to select slice for UE. The message includes requested NSSAI, subscribed NSSAI, UE location.
  5. NSSF determines allowed NSSAI and selects target AMF based on information provided by AMF. NSSF sends Slice Selection Response to AMF. The message includes allowed NSSAI, mapping between the S-NSSAI in the
  Allowed NSSAI and the NRF and AMF list. The AMF list may include AMF IP address list or FQDN list. AMF determines whether it is the target AMF based on whether it is included in the AMF list. If it is the target AMF, then the steps in box (B) are performed, the steps in box (A) are skipped. Otherwise the steps in box (A) are
  performed, the steps in box (B) are skipped.
  6. AMF determines that it is not the target AMF, it sends Re-route NAS message to RAN. The message includes the Registration Request and the identifier of the target AMF.
  7. RAN sends the Registration Request to the target AMF.
  8. Authentication and security are performed.
  9. Target AMF sends Registration Accept message to UE.
  10. AMF determines that it is the target AMF. Authentication and security are performed.
  11. AMF sends Registration Accept message to UE.

# Core Network Slice for Resources

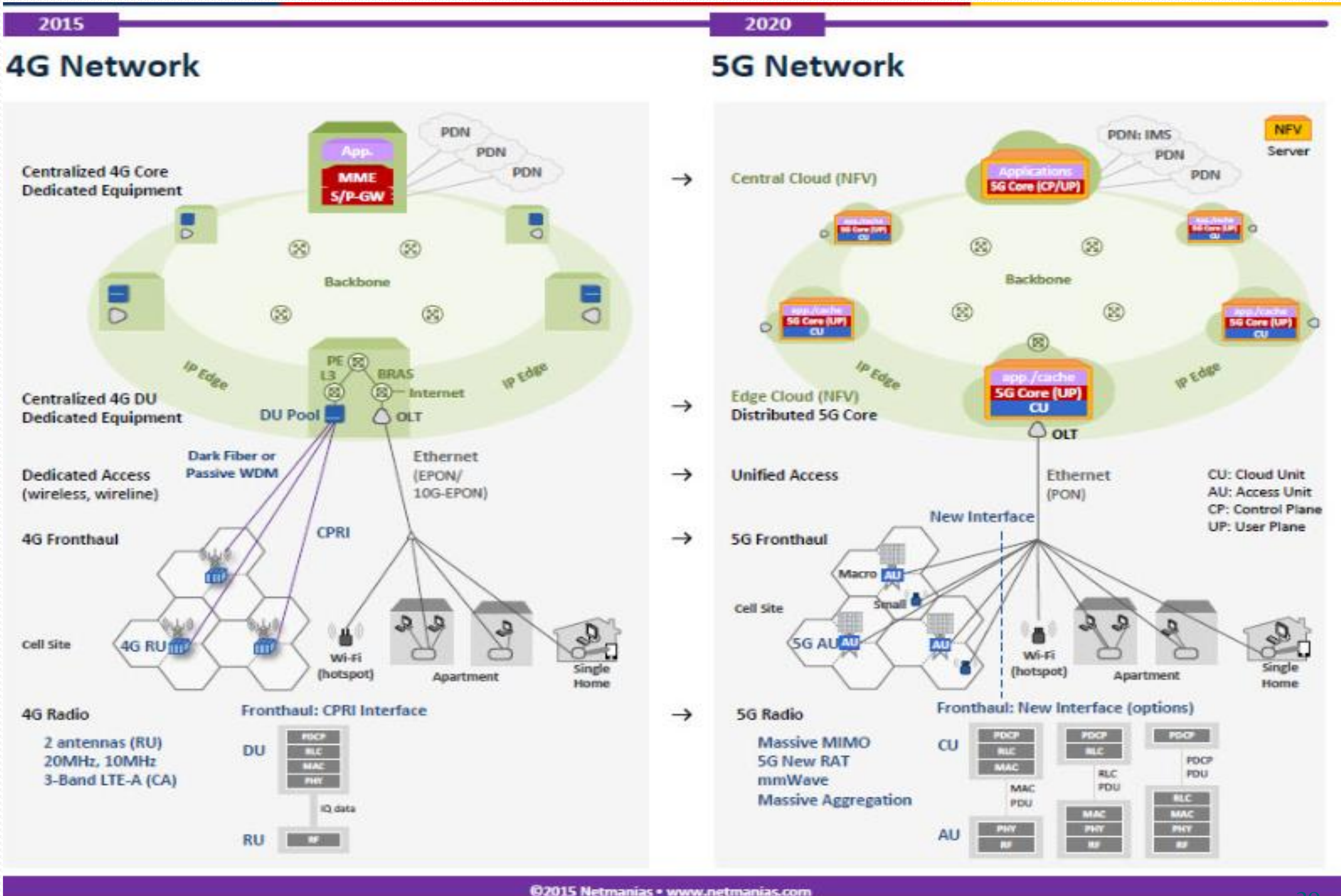

UE/Device may be served by one/more CN slices according to the services/ Requirements

# Transport Slicing




- **Physical Slicing**: slicing of the network resource (e.g., assignments for network elements, ports, links, racks etc.)
- **Hard Pipe Slicing**: L1 channel slicing (e.g., the channels of WDM, OTN, FlexE)
- **Soft Pipe Slicing**: Ethernet / IP virtual network segment (e.g., L2 VPN, L3 VPN, L2 / L3 Bridge)
- **Service Slicing**: service type slicing (e.g., URLLC, eMBB, mMTC)

# RAN Slicing: 4G to 5G Network Architecture



# Network Slicing Requirements

## Transport, Core Network, and RAN

### Platform Resources

1. Time-Frequency
2. vCPU, vRAM, vStorage

### Control of Platform Resources

1. Resource to Slice – Static / Dynamic
2. Control / User Plane separation

### Algorithms

1. Scheduling
2. Handover Management
3. Admission Control
4. SON
5. QoS shaping
6. Security

### Slice Management

1. User Management
2. Flow Management
3. SDN Control
4. Service Orchestration

# Few Research Steps in Network Slicing @ IIITB

- Radio and Platform Resource Management at RAN using Contention and Non-contention
- Network Slice Profile characteristics and Profile Migration
- SDN control of RAN slices for Control and User Plane
- Slice-specific admission control
- Network Slice Backhaul Scheduling and Management

# Research on Network Slicing – IIITB

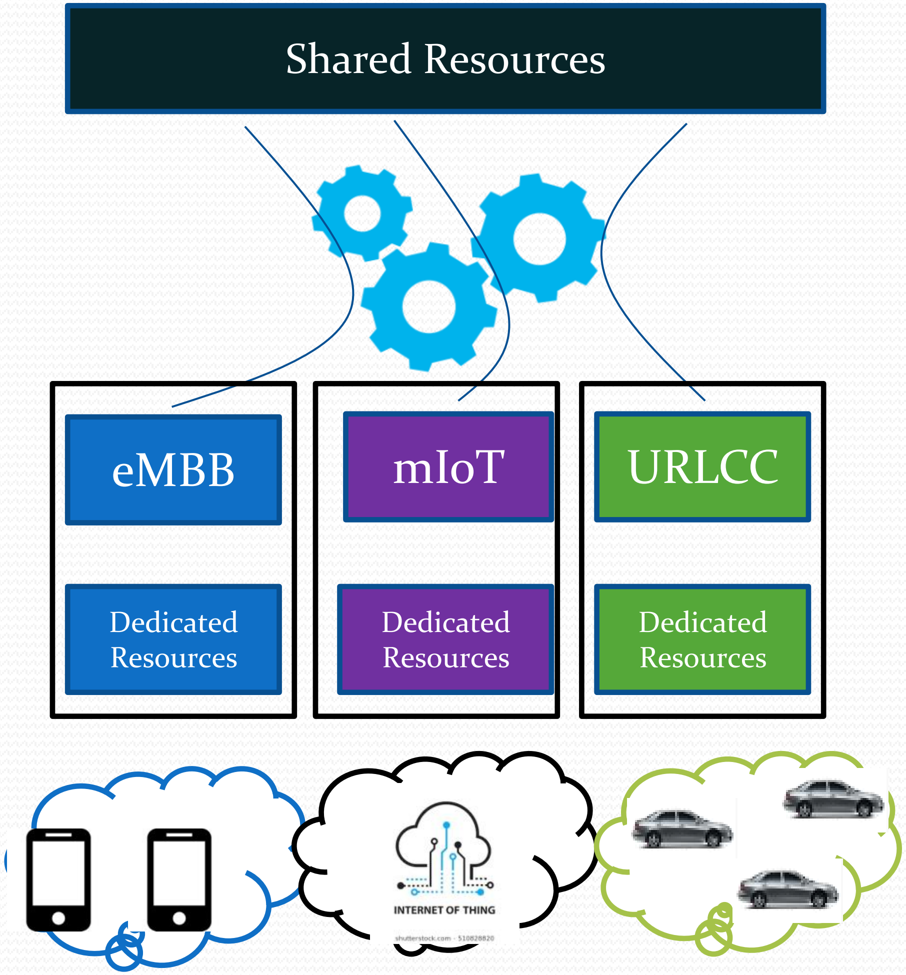
## Radio  and Platform Resource Management at RAN

- Complete system resources (Radio, Platform etc.,) of RAN is divided across statically or dynamically for the different Network Slices

- Dynamic resource allocation take time study and hence need longer time → how to reduce it or go back to static allocation?
- Controlling the association / de-association of Users switch multiple Network Slices and its resources without contention

- Network Slice scheduler monitors the resource usage within the Slices and optimally demands system resources in a contention manner

- Study the Network Performance Parameters with respect of QoS



# Research on Network Slicing – IIITB

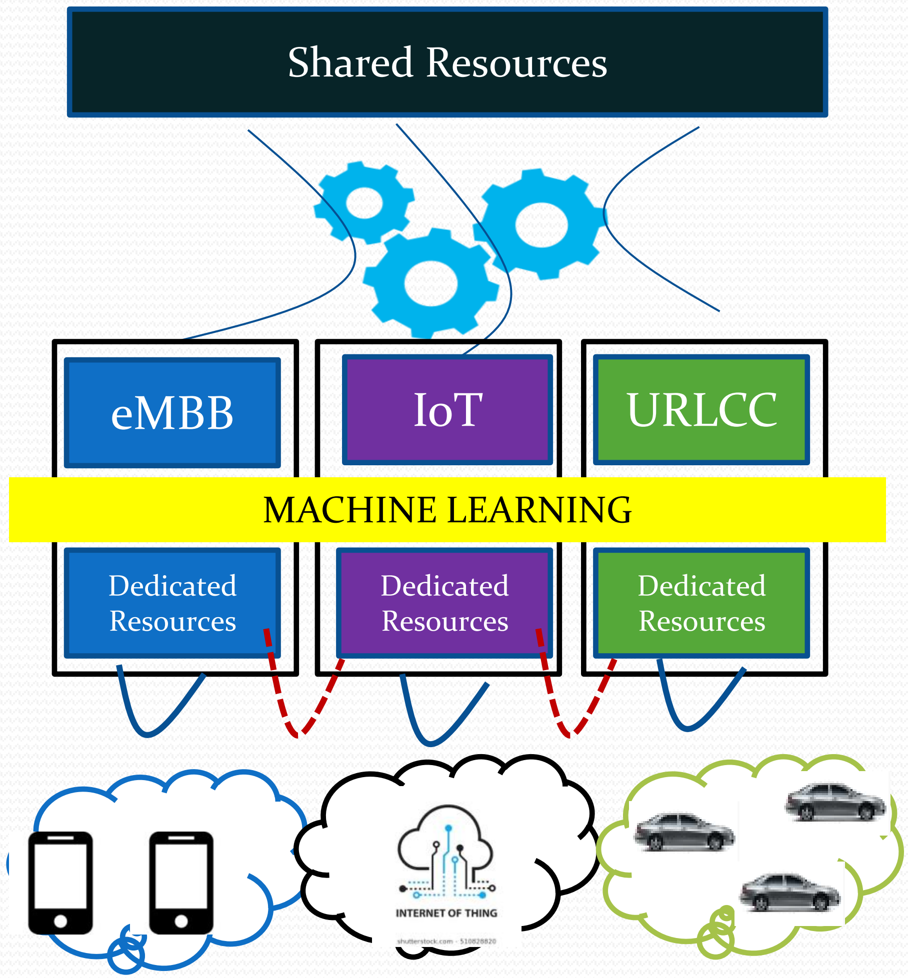
## Network Slice Profile characteristics and Profile Migration at RAN

- Creation of Network Slice Sub-Profile based on additional multiple criteria like User mobility, Channel Condition, Round trip time, Latency characteristics, User Throughput, Burst traffic characteristics, block error rate, round trip latency, packet drop rate, channel quality variation handling, random access attempt, retransmission etc. etc.,

- Network Slice Resource usage model creation and future utilization prediction using latest machine learning concept

- Network Resource migration across Slices based on Profile usage model and operational characteristics

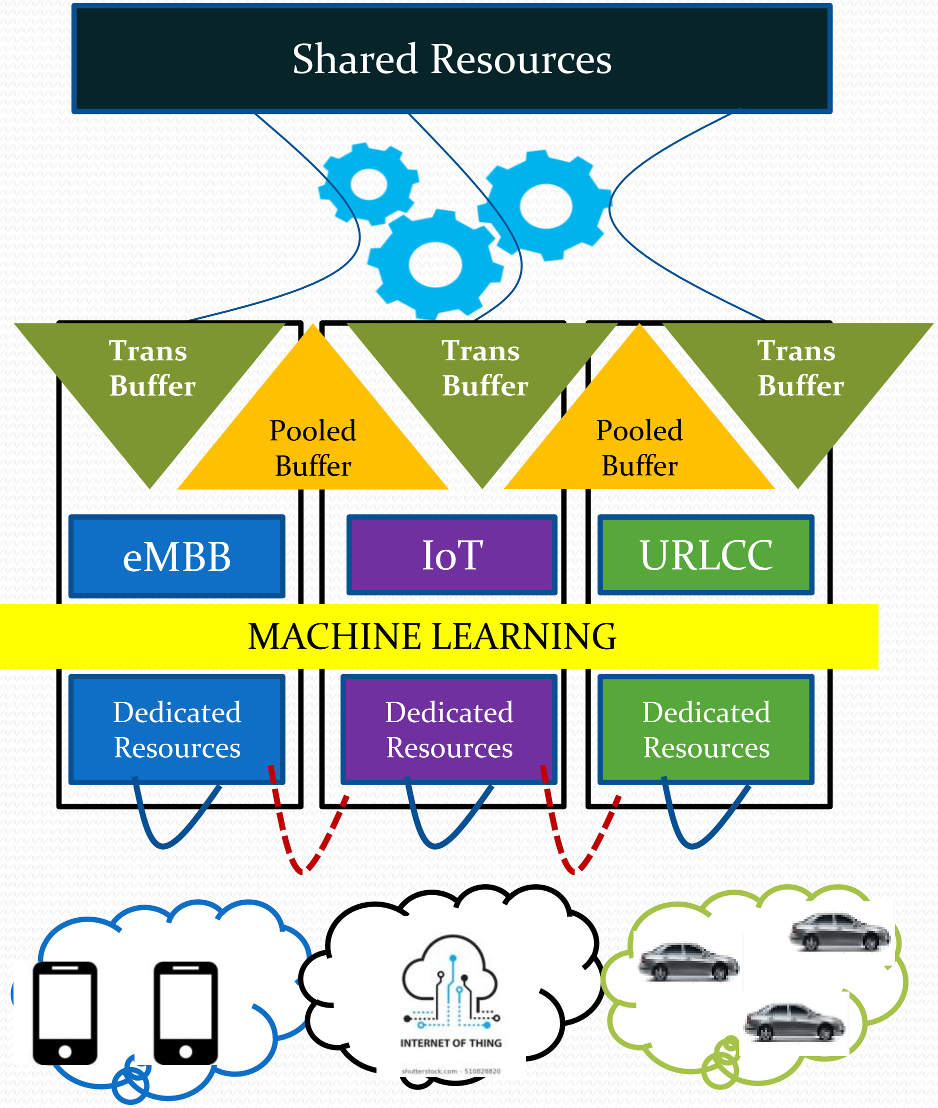
- Study the Network Performance Parameters with respect of QoS



# Research on Network Slicing – IIITB

Slice-specific admission control at RAN

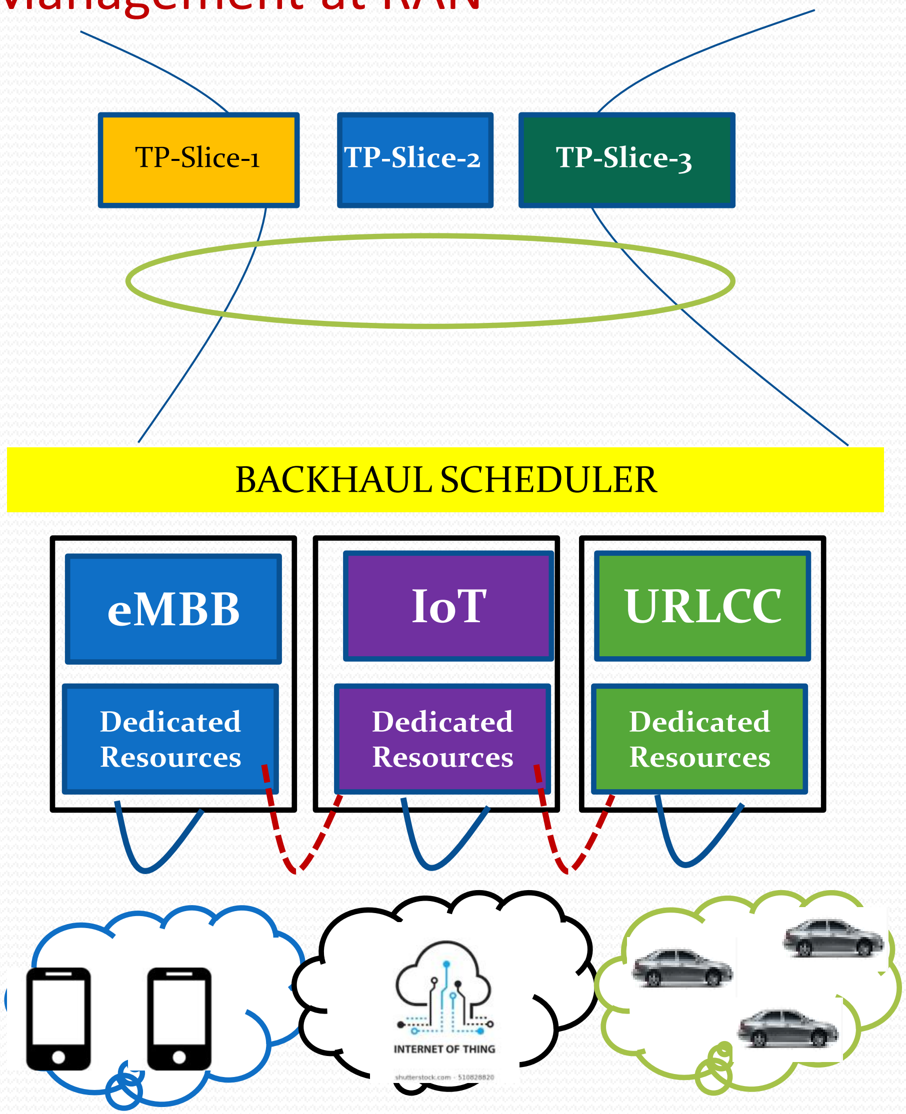
- Controlling the number of users admitted per Network Slices
- Using machine learning, analyze the Network Slice Profile current operational usage and future usage
- Create of dynamic buffer and User Slice admission based on Profile future usage, Pre-emption etc.,
- Study the Network Performance Parameters with respect of QoS



# Research on Network Slicing – IIITB

## Network Slice Backhaul Scheduling and Management at RAN

- Static and Dynamic mapping of the Network Slices with the Transport Slices
- Controlling and Managing the packets to/from Backhaul from/to Network Slices, Link Aggregation
- Offload control for Local Breakout
- Study the Network Performance Parameters with respect of QoS



# Summary: Benefits of Network Slicing

- **Device capability based access**
  - A device that is simple and requires only access to a single slice. e.g., a water meter reader can find a radio interface that is tailored to very small, infrequent messages RAN attaches the device to a simple, efficient authentication process connected to a single core water meter network slice
  - A device that is more complex and provides rich functionality, such as a smartphone. E.g. device can be authenticated and attached to a diverse set of network slices that are each tailored to a specific purpose: streaming video, voice calls, internet browsing, chatting
- **Spawning of specific Network Functions**
- **Guaranteeing E2E QoS**
- **Mobile edge computing /edge cloud processing**

# Summary (Continue)

- In 5G era, the communication will further change our society by providing solid foundations to realize the "Internet of Everything":
- To enable such a vision it is essential to have a deep understanding of the various requirements of the vertical industries and serve them without sacrificing efficiency while keeping costs low.
- **As one of the key features for 5G, network slicing allows a flexible system** architecture able to deal with those requirements.
- 5G is a communication vision turning point, which will have deep impacts on telecommunication as well as vertical industries.

# Thank You!

# References

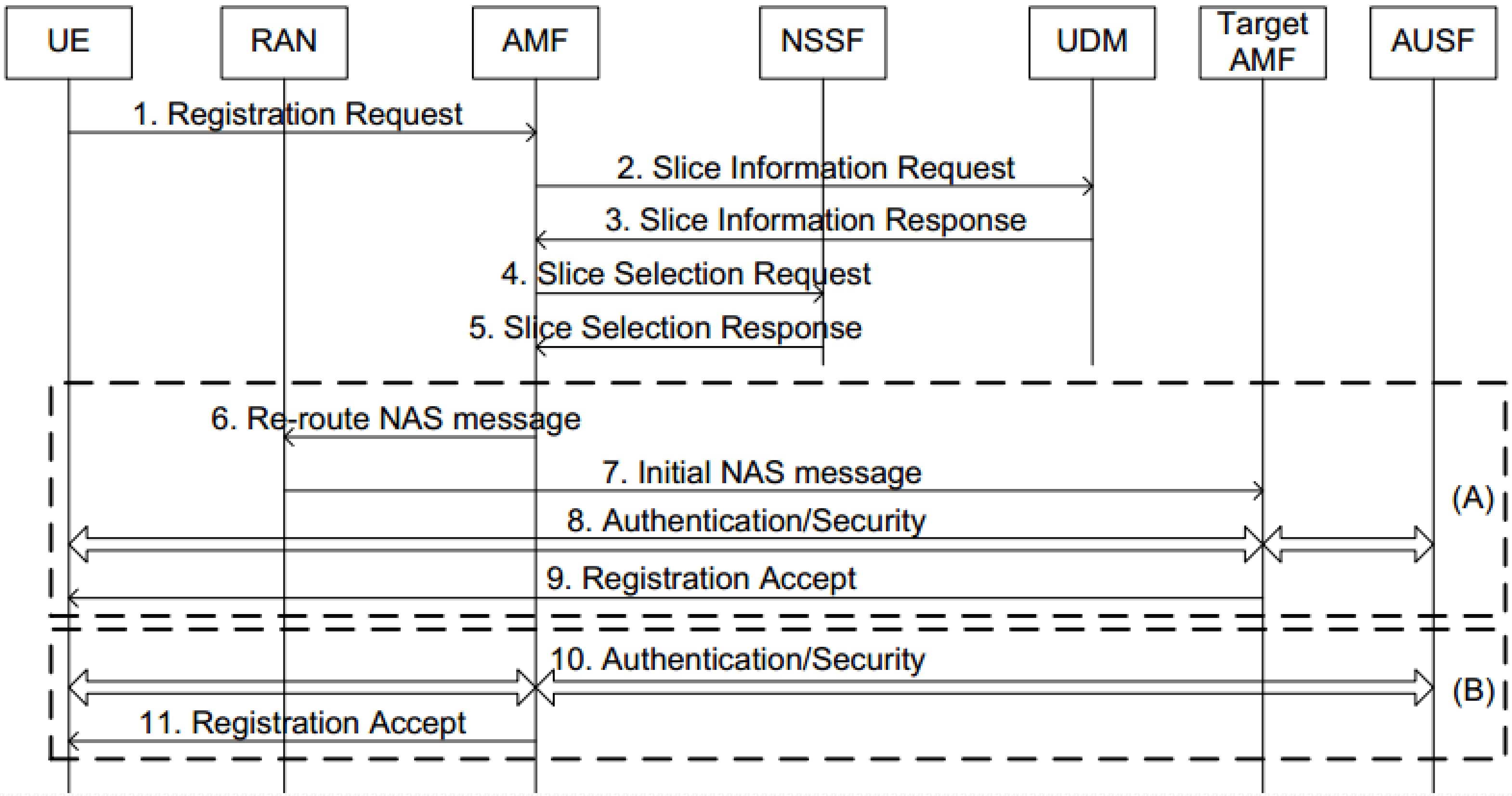
# Extra Slides

# Silo Architecture of Telecom

- In past, every telecom operator had their own network deployment → proprietary hardware and software systems
  - Example: POTS based Telephone system (GSM, IS95A/B
- Each of these telecom network deployment did not have any common Protocol Layers → physical, Layer2, Layer-3 different
- Interworking among the telecom operator was difficult
- Each of the telecom operator had their own communication devices enabled with their specific application (e.g. voice) with specific use case deployment
- Later part of the technological development, the interworking of the application was done using the **Gateway** products.

- Authentication Server Function (AUSF): supports authentication server function.
  • Access and Mobility Management Function (AMF): access control, mobility control, transparent proxy for routing
  SM message.
  • Unified Data Repository (UDR): storage and retrieval of data by the UDM, PCF or NEF.
  • Unstructured Data Storage Network Function (UDSF): storage and retrieval of information as unstructured data by
  any NF.
  • Network Exposure Function (NEF): expose the services and capabilities provided by 3GPP network functions.
  • NF Repository Function (NRF): maintains NF profile, supports service discovery.
  • Policy Control function (PCF): decides the policy and provides them to the control plane function.
  • SMF (Session Management Function): manages the PDU session e.g. PDU session establishment, modify and
  release.
  • Unified Data Management (UDM): authentication credential processing, access authorization, registration/ mobility
  management and subscription management.
  • User plane Function (UPF): handles the user plane traffic, e.g. traffic routing & forwarding, traffic inspection and
  usage reporting, handling.
  • Application Function (AF): interacts with the 3GPP Core Network (CN) to provide services.
  • Network Slice Selection Function (NSSF): selects the NSI, determines the allowed network slice selection
  assistance information (NSSAI) and AMF set to serve the UE.

# Registration Procedure in Core N/W

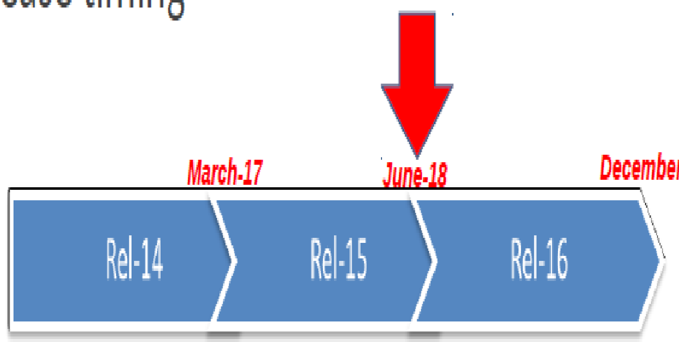


# Steps of Registration in Previous Slides

- 1. UE sends Registration Request to AMF. The message includes requested NSSAI.
  2. AMF sget subscription data from UDM.
  3. UDM responds subscription data to AMF. The subscribed NSSAI includes subscribed S-NSSAI.
  4. AMF sends Slice Selection Request to NSSF to select slice for UE. The message includes requested NSSAI,
  subscribed NSSAI, UE location.
  5. NSSF determines allowed NSSAI and selects target AMF based on information provided by AMF. NSSF sends
  Slice Selection Response to AMF. The message includes allowed NSSAI， mapping between the S-NSSAI in the
  Allowed NSSAI and the NRF and AMF list. The AMF list may include AMF IP address list or FQDN list.
  AMF determines whether it is the target AMF based on whether it is included in the AMF list. If it is the target
  AMF, then the steps in box (B) are performed, the steps in box (A) are skipped. Otherwise the steps in box (A) are
  performed, the steps in box (B) are skipped.
  6. AMF determines that it is not the target AMF, it sends Re-route NAS message to RAN. The message includes the
  Registration Request and the identifier of the target AMF.
  7. RAN sends the Registration Request to the target AMF.
  8. Authentication and security are performed.
  9. Target AMF sends Registration Accept message to UE.
  10. AMF determines that it is the target AMF. Authentication and security are performed.
  11. AMF sends Registration Accept message to UE.

# 5G Timeline

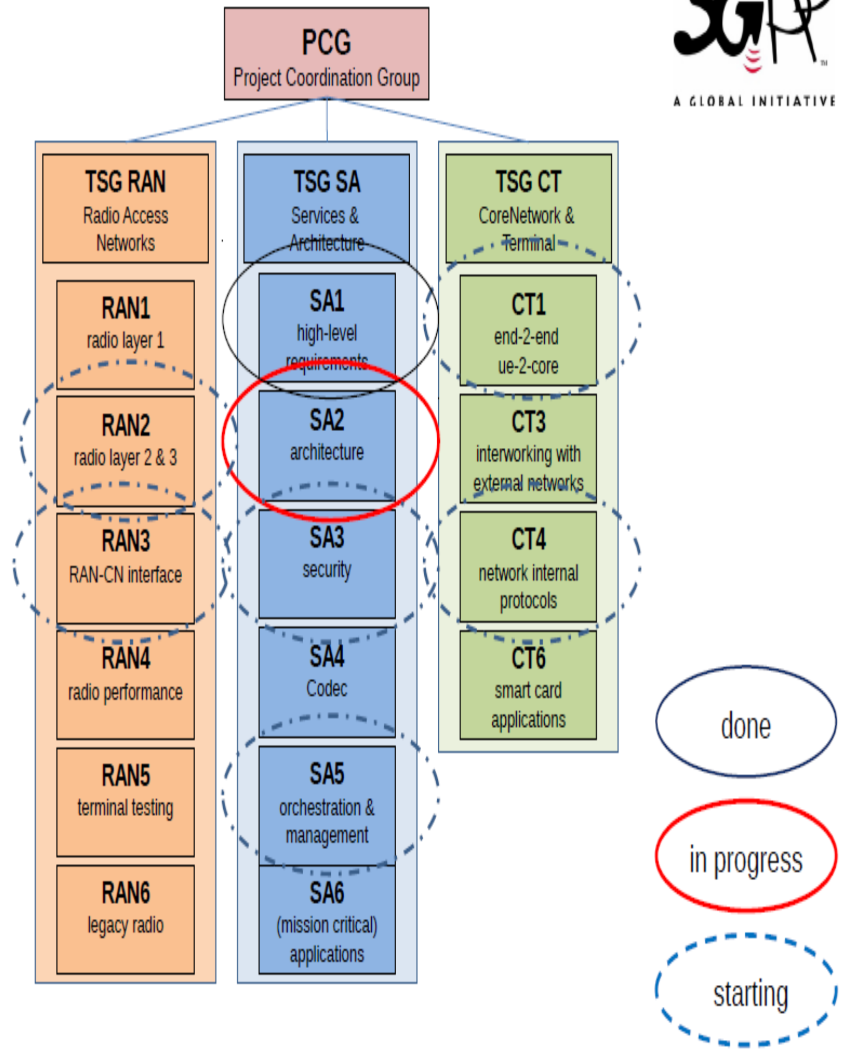
## Release timing




## Two phases for the normative 5G work

- Phase 1 (Rel-15) to be completed by June 2018
  addresses the more urgent subset for commercial deployments

- Phase 2 (Rel-16) to be completed by March 2020
  IMT 2020 submission, addresses all identified use cases & requirements




3GPP A GLOBAL INITIATIVE

# 5G Network Slicing in SA2

- Network Slice
  - A logical end-to-end network
  - Dynamically created
- Different slices for different services types
  - Committed services – slice types
  - Dedicated customers
- Comprise of
  - 5G Core Network (CP (Control Plane) & UP (User plane))
  - 5G Radio Access Network
  - Interworking Functions to non-3GPP Access Networks
- UE connects
  - Max 8 slices in parallel (till recent standard)
  - Common AMF (Access and Mobility Function) for one UE in all slices

# 5G Network Slicing in SA5

- SA5 identified related use cases, potential requirements and potential
  - lifecycle management
  - fault management
  - configuration management
  - performance management
  - policy management

- SA5 has identified and describes the following management functions
  - the Communication Service Management Function
  - the Network Slice Management Function
  - the Network Subnet Slice Management Function

- Upcoming work are:
  - enable the provisioning of network slices for 5G networks and services
  - information model to support different Slice/Service type
  - network slice template and define the relation of network slice template and
  - network slice instance
  - requirements for transport and virtualization

# Network Slicing – Parameters

- S-NSSAI – single network slice selection assistance information
  - SST – slice type, describes expected network behavior
  - SD – slice differentiator, optional, further differentiation
- S-NSSAI can have standard or network-specific values
  - Standard SST values: *eMBB, URLCC, MIoT*
- NSSAI is a collection of max 8 S-NSSAI
- UE sends NSSAI – based on which related slice(s) are selected

| Slice/Service type | SST value | Characteristics. |
|---|---|---|
| **eMBB (enhanced Mobile Broadband)** | 1 | Slice suitable for the handling of 5G enhanced Mobile broadband, useful, but not limited to the **general consumer space** mobile broadband applications including<br>- streaming of High Quality Video,<br>-Fast large file transfers etc.<br><br>It is expected this SST to aim at supporting High data rates and high traffic densities |
| **URLLC (ultra- reliable low latency communications)** | 2 | Supporting ultra-reliable low latency communications for applications including,<br>- industrial automation,<br>- (remote) control systems. |
| **MIoT (massive IoT)** | 3 | Allowing the support of a large number and high density of IoT devices efficiently and cost effectively. |

**TS 23.501, section 5.15.2.2-1**

# Broad Network Slicing Architecture [3]

- Network Slicing has major three layers
  - Infrastructure Layer
  - Network Slice Layer
  - Network Management Layer
- Infrastructure Layer: Providing actual and virtual resources to slice → computing, storage, connectivity, etc.
- Network Slice Layer: runs over infrastructure layer; provides network functions to form End-to-End (E2E) logical network
- Management Layer: designed to mange network slices
  - Business Support System (BSS)/operation support system (OSS)
  - Network Slice Management (NSM) System
  - Ensures Service Level Agreement (SLA) requirements