Enabling Cross-layer Receiver Oriented Dynamic Multicast in Cellular Access

Ravi Ravindran, Parishad Karimi

(ravi.ravindran@Huawei.com/Parishad.Karimi@Winlab.com)

IETF/ICNRG-101, March, 2018

Agenda

- Motivation
- LTE-B/E-MBMS System
- ICN Enabled Dynamic Multicast
- Why ICN helps ?
- ICN-MAC for Cellular Access
- DL Scheduling Solutions
    - Centralized and Distributed Approach
- Opportunities and Challenges

Motivation

- Just as in the core, broadcast/multicast in wireless access improve spectrum utilization significantly.

- Spectrum takes the significant part of a cellular access operator's CAPEX.
    - $20B worth spectrum auctioned by US FCC in 2017, this is just the beginning
        - https://www.wirelessdesignmag.com/blog/2017/04/fcc-spectrum-auction-helps-companies-prepare-5g

- Any amount of saving in the radio access is welcome
    - Dinesh Bharadia et al "**Full Duplex Radio**", Sigcomm, 2013 – PHY layer solution.
        - Led to **Kumu Networks**

→ If ICN can help save radio spectrum resources, it would be more compelling than the bandwidth saving in the infrastructure.

Cellular Broadcast

- **No notion of multicast on the radio interface**
  - Only Unicast or Broadcast
  - Resources reserved for the duration of the event
- **Broadcast good for pre-planned events (e.g. Mobile TV, Live events)**
- **Several solutions for media broadcast have had limited commercial success**
  - DVB-H (Europe)
  - Media-FLO (US)
  - ISDB-T (Japan)
  - T-DMB
  - ..
  - LTE-B

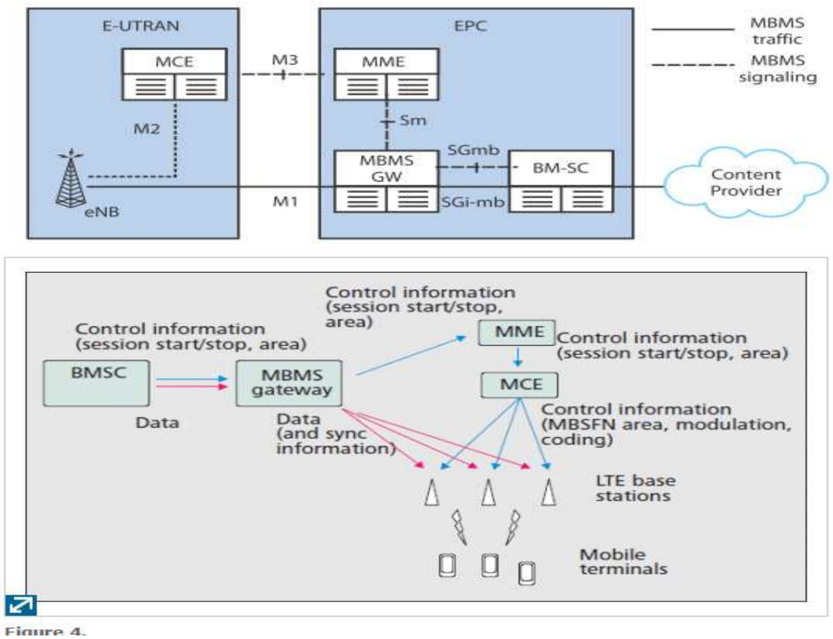
[1] 5G X-Cast white paper, "Broadcast and Multicast Communication Enablers for the Fifth-Generation of Wireless Systems", Part of the 5GPPP Project, Nov 2017

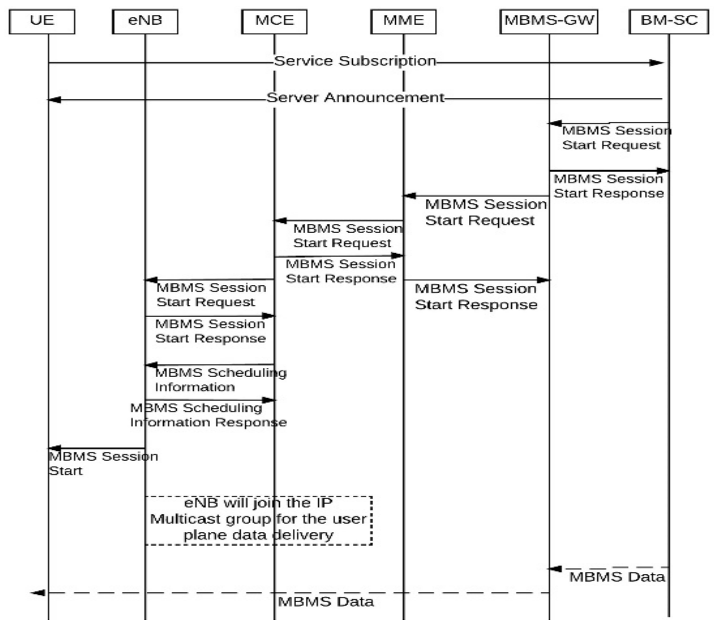
LTE-B

- **Towards this, E-MBMS (LTE-B) extensions were proposed to enable multicasting in LTE, since Rel-6**
  - Applications included live TV to general audience or live events, e.g. scoped within a stadium

- **LTE-B is not a commercial success because of many reasons**
  - User behavior moving to VoD based services
  - Pre-provisioned broadcast spectrum resources under-utilized.
  - One way broadcast, without any feedback capability from UE
    - Adapts to the weakest receiver because of lack of feedback
  - Latency in setting up multicast sessions because of control and data channel setup latency
  - No IP multicast in the core (e.g. EPC), still difficult to adapt to on-demand content delivery
  - Scheduling challenges, i.e choosing the right MCS that meets diverse user requirements, e.g. cell edge vs mobile vs static users
  - New hardware capability because of new UP/CP
  - Lack of standard API between operators and the content providers

E-MBMS Overview





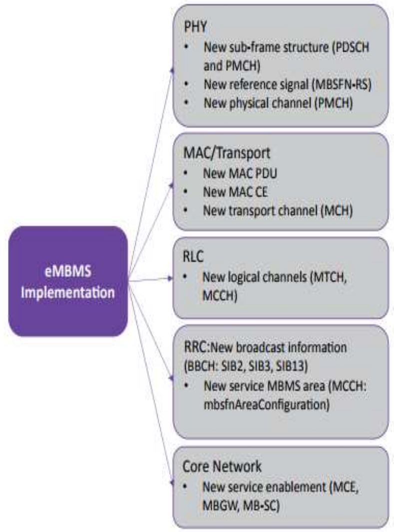


Fig 1: LTE MBMS Architecture

- EMBMS employs specialized Control/User plane for multicast delivery.
- This includes the DP interaction between MBMS-GW that IP multicasts the data to a set of eNodeB, then it signals MCE for allocating appropriate resources at the eNodeB through the MME.
- The delay between the control and user play data delivery is in order of ~10sec [1], which wouldn't work for VoD or real-time applications like conferencing.

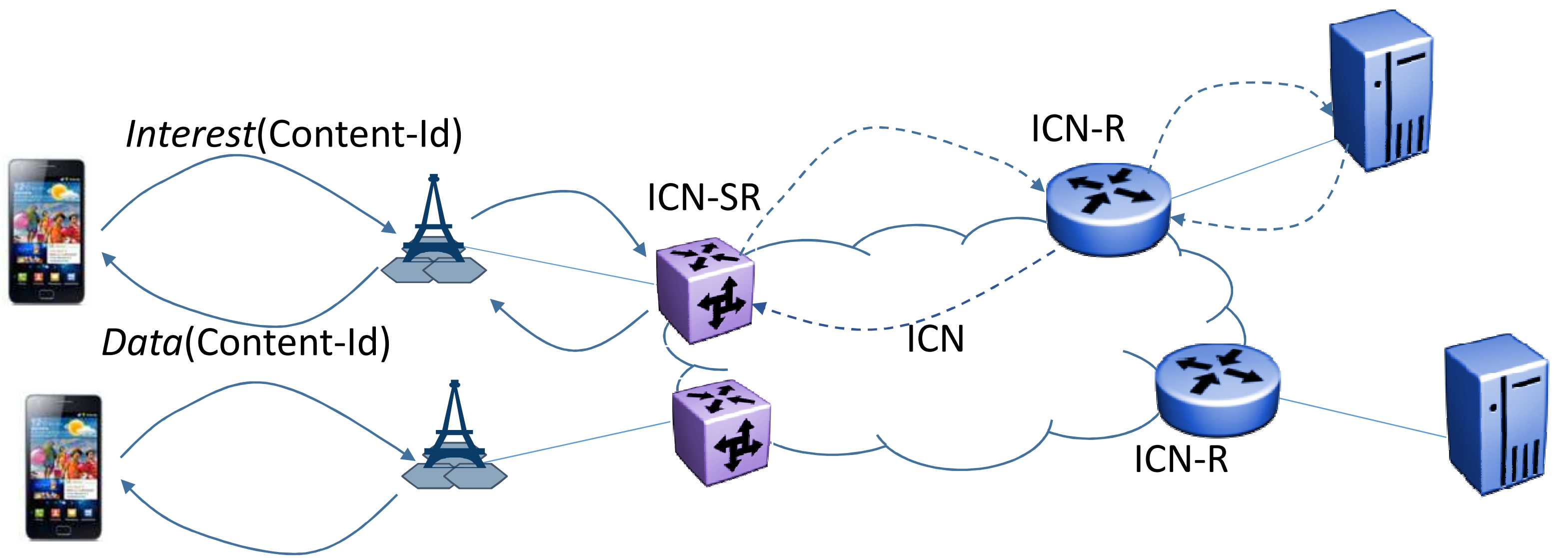
[1] David Lecompte et al, "Evolved Multimedia Broadcast/Multicast Service in LTE Advanced: Overview and Rel-11 Enhancements", IEEE Comm. Magazine, 2012

Single Frequency Networks in the Access

- E-MBMS uses SFN broadcasting in the access
- Multiple eNodB's are synchronized to broadcast the same content over the same channel, hence called Single Frequency Network
- The set of participating eNodeB form an MBMS-SFN
- This improves the quality of reception at the UE.
- However, this overall leads to poor resource utilization across multiple cells as stated earlier, if there is lack of consumers for the service.

ICN enabled Dynamic Multicast in Wireless Access

- ICN enables dynamic multicast in the core network, but the access is mostly based on unicast proposals today.
- Work on ICN over Ethernet/Wifi/802.11p link layers.
- We propose mechanisms to exploit ICN multicast in the wireless cellular access.



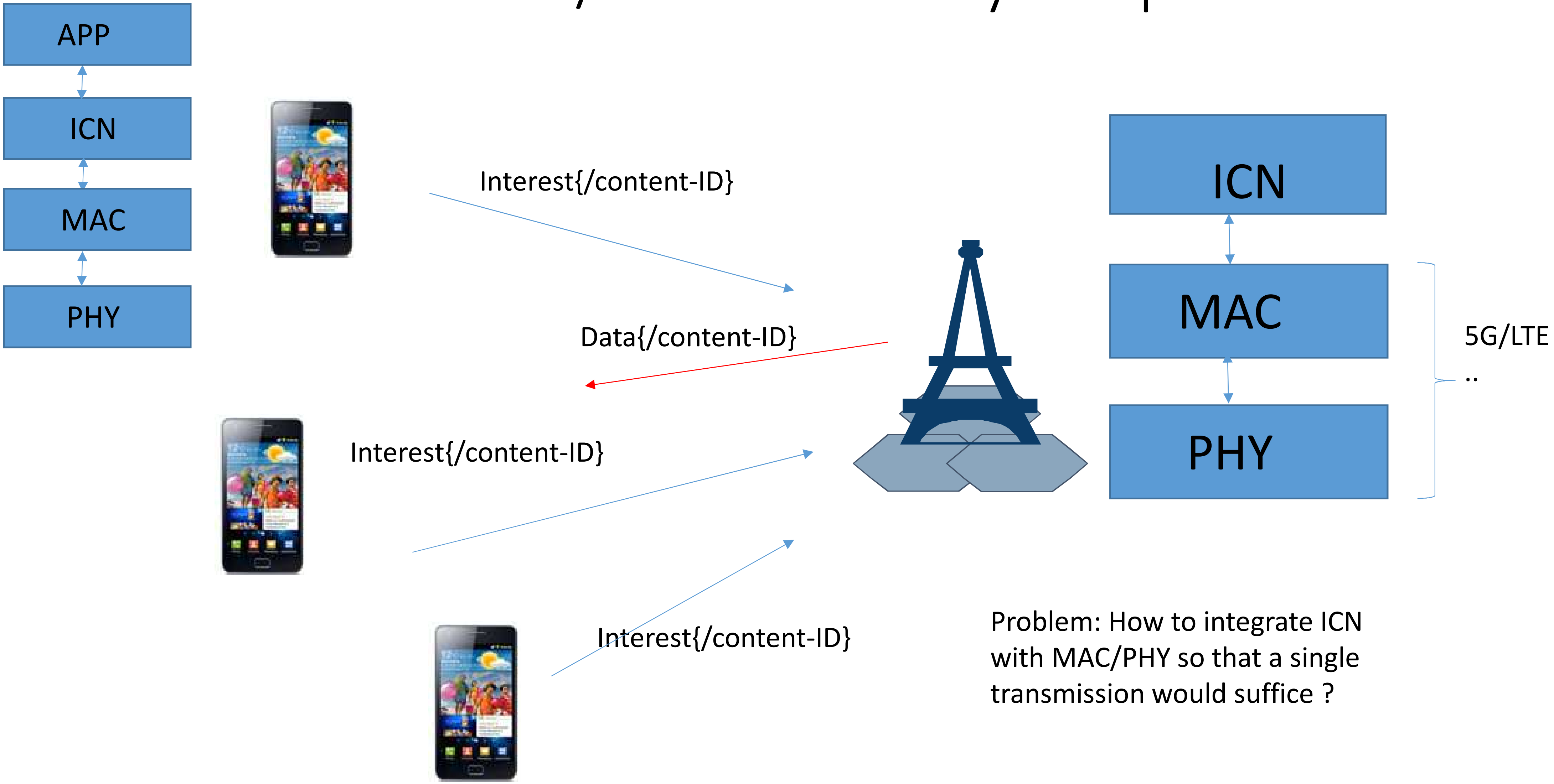
Why ICN helps in cellular access multicast?

- ICN allows dynamic demand adaptation to aid content distribution, enabling popular content caching closer to users, e.g. eNodeB
- Removes host level notion, MAC can operate in an information-centric manner too.
- Unicast and Multicast becomes indistinguishable.
- Allows cross layer, i.e. ICN and MAC/PHY layers, to improve efficiency
- ICN Interests can carry wireless parameters in its requests, e.g. CQI allows cross layer adaptation
  - Interests/Data offers a bi-directional paradigm that can assist the MAC layer.

ICN-based Multicast vs. eMBMS

	ICN-based Multicast	**eMBMS**
Multicast setup and management	Low-overhead and low-latency	High signaling overhead and specialized gateways
	Self organizing operation	Always initiated and maintained by server and gateway entities
QoS incorporation	Supported by exploiting ICN semantics in MAC layer	Not Supported for fine-grained control, only focused on MTC
Mobility and dynamic leave/join	Handled at short time-scales	Not handled
Content delivery	Dynamic popularity-based edge-caching of content (at the eNBs)	Cross-layer optimization for content delivery not supported

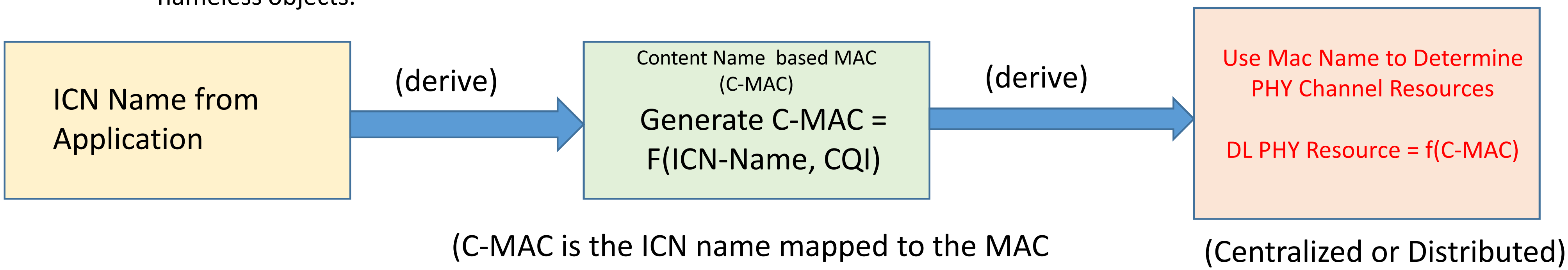
ICN and MAC/PHY Cross-Layer Optimization



Problem: How to integrate ICN with MAC/PHY so that a single transmission would suffice ?

Basic Idea – Making MAC operate on dynamic Multicast ID

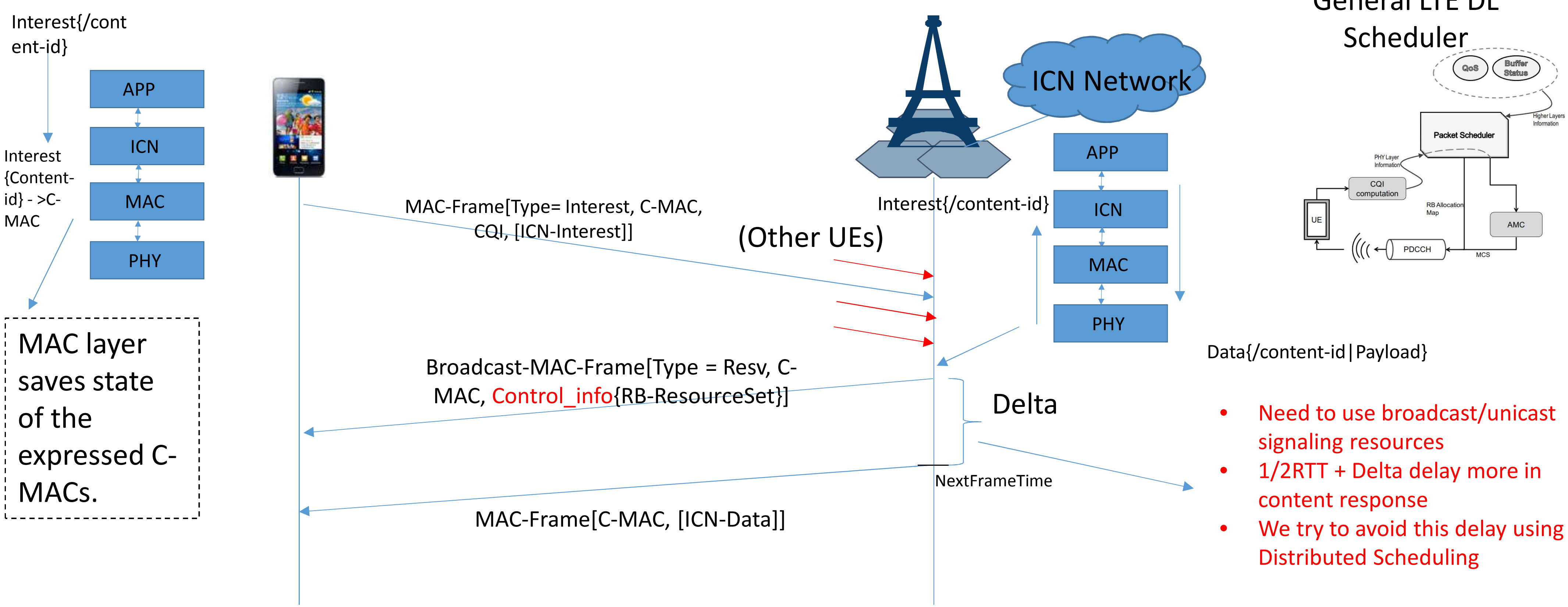
- Make Unicast and Multicast share the same control and data channel without any special functions to support multicast.
- We enable C-MAC to be derived from content name, hence allowing the MAC layer to operate in an ICN manner
  - Multiple approaches to derive C-MAC from ICN name, it can come from the applications, e.g. from manifest or generated locally using a hash function
  - We deal with hash collisions issues later in the presentation. Generally hash collision is an ICN problem, for e.g.in the use of nameless objects.



(C-MAC is the ICN name mapped to the MAC layer. The ICN name in the UE is mapped to a dynamic multicast ID. This enables MAC to operate in an ICN manner, i.e. host independent, enabling one time Content Push in the Air Interface)

Centralized eNodeB Driven Scheduling

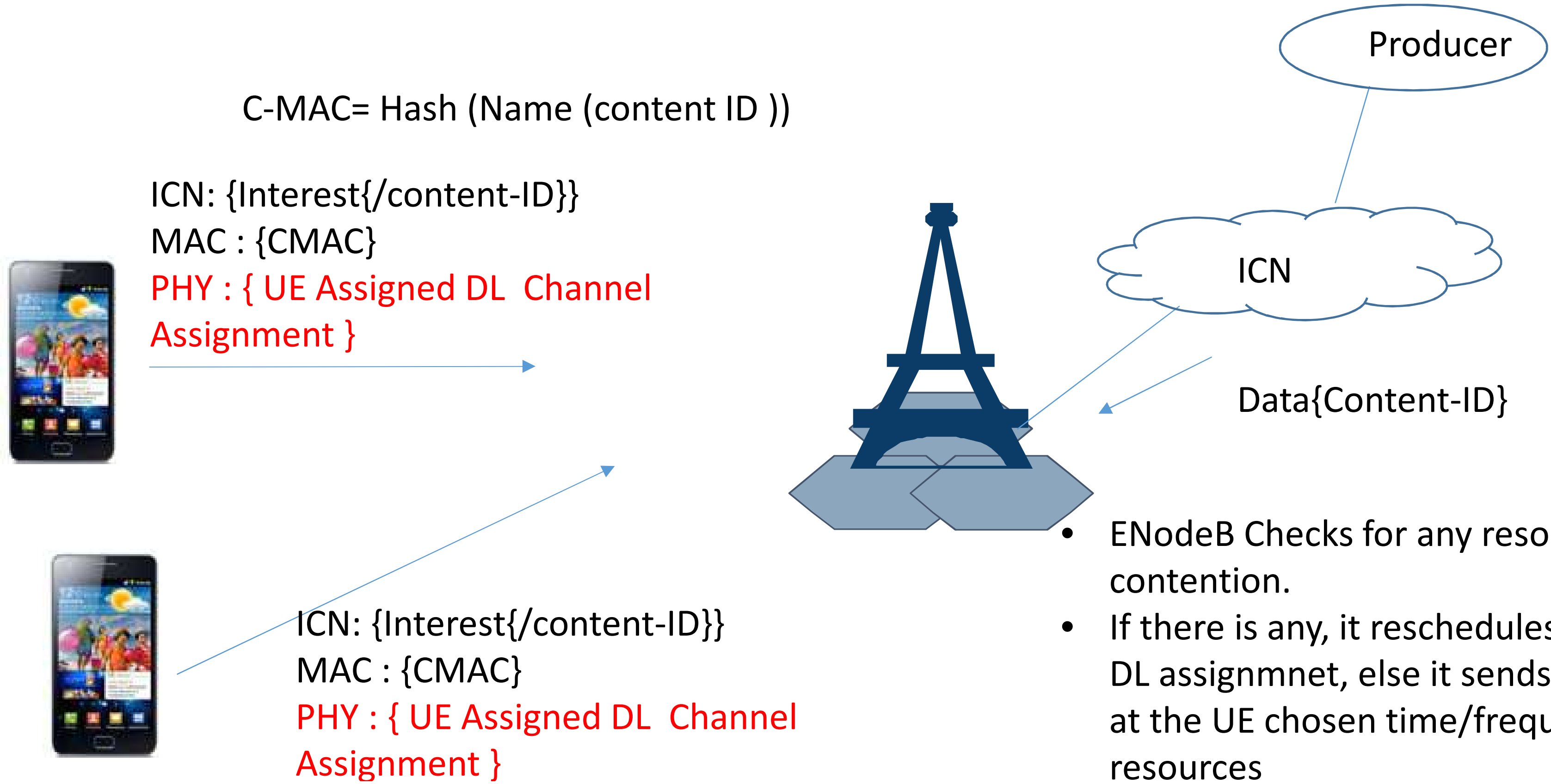
- One approach is to allow centralized eNodeB scheduling based on C-MAC (shared by multiple UEs).
- Like today UEs have to be updated about the PHY channel allocations and then receive multicast content.
- **For this broadcast/unicast control channel resources has to be reserved a priori, leading to wastage of resources. Also the broadcast forces all UE to process information that may not be relevant to them.**
- This mode of scheduling is also not efficient considering higher RTT for the content response



- Need to use broadcast/unicast signaling resources
- 1/2RTT + Delta delay more in content response
- We try to avoid this delay using Distributed Scheduling

UE Driven Scheduling

- Here we distribute scheduling task between UE and the eNodeB. UE first chooses the RBs it decides to listen on in the DL. Any contention is then resolved by the eNodeB.

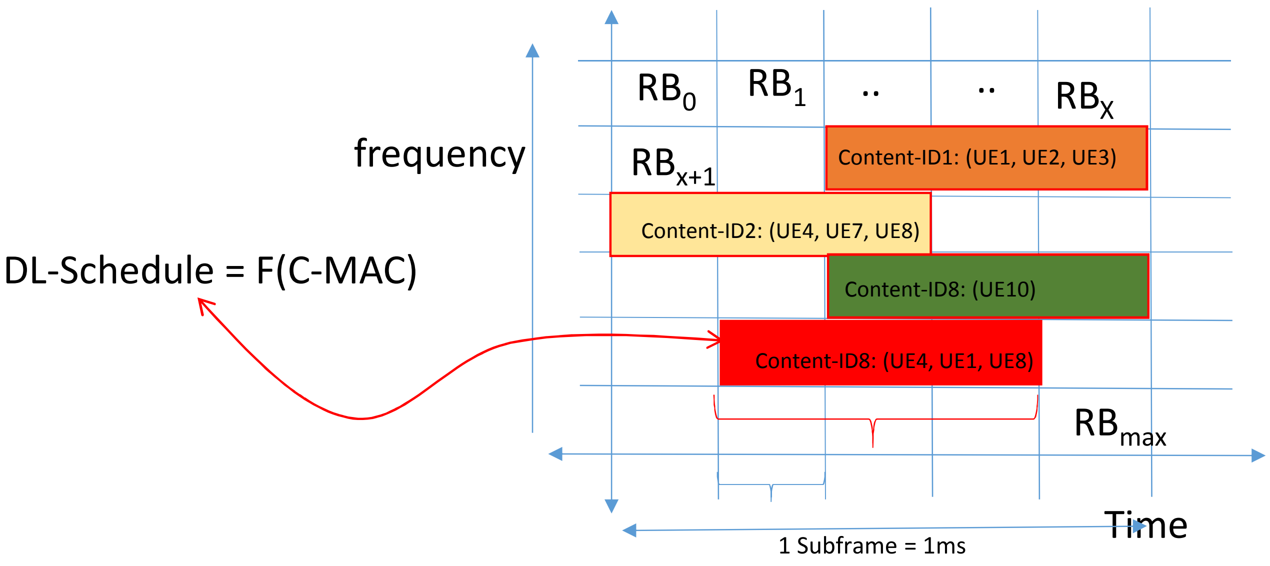


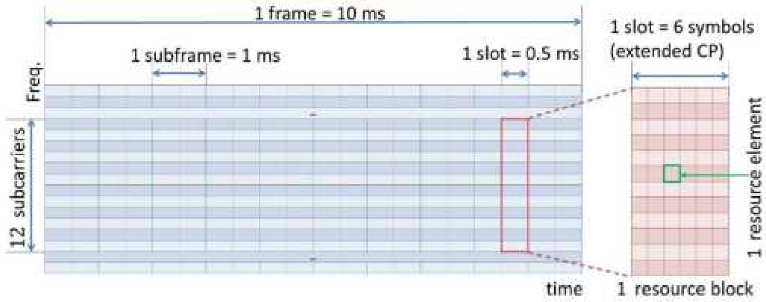
- ENodeB Checks for any resource contention.
- If there is any, it reschedules the DL assignmnet, else it sends dat at the UE chosen time/frequenc resources

UE Driven Scheduling: Handling DL Channel Assignment

- The goal is to use distributed scheduling, moving away from traditional centralized scheduling to handle dynamic multicast, so that all UEs asking for the same content in a Sub-frame time window (e.g. 1ms for LTE) can scheduled on a single transmission slot.
- The idea is to use C-MAC to derive DL scheduling resources (DL-Schedule).

DL Resource Matrix





Example LTE frame

UE Driven Scheduling: Algorithm to Map C-MAC to RB ID

- The objective here is to map with minimum collision at the DL PHY resource level using contextual C-MAC.
- We assume enumeration of the PHY channel as {RB_0, $RB_{1\ldots}$,RB_{MAX}}., e.g. 1RB = 180Khz X 1 Slot
  - Instead we could also work at the level of a Resource Element RE, in LTE, a RB = 12 sub-carrier X 1 Slot, where 1 Slot = 0.5ms, I Slot = 6 Symbols (~50micro-s), hence , 1RB = 84 or 72 Res
  - For 1GHZ spectrum MAX_RB could be 5000RB, or 500,000 REs, this is function of Subcarrier spacing and Symbol duration
- 5G-NR Numerology is similar to LTE, but with more flexibility adapting to different application
- We propose the following Algorithm

UE Driven Scheduling: Algorithm to Map C-MAC to RB ID

S1: Determine a DL-Schedule based on the C-MAC and maximum number RBs available for DL scheduling. Assuming C-MAC as a random ID, *DL-Schedule(C-MAC) = f(C-MAC, RB_{MAX})*

S2: Determine the number of RB based on the estimation of the expected MTU (from ICN layer) and the CQI of the user.

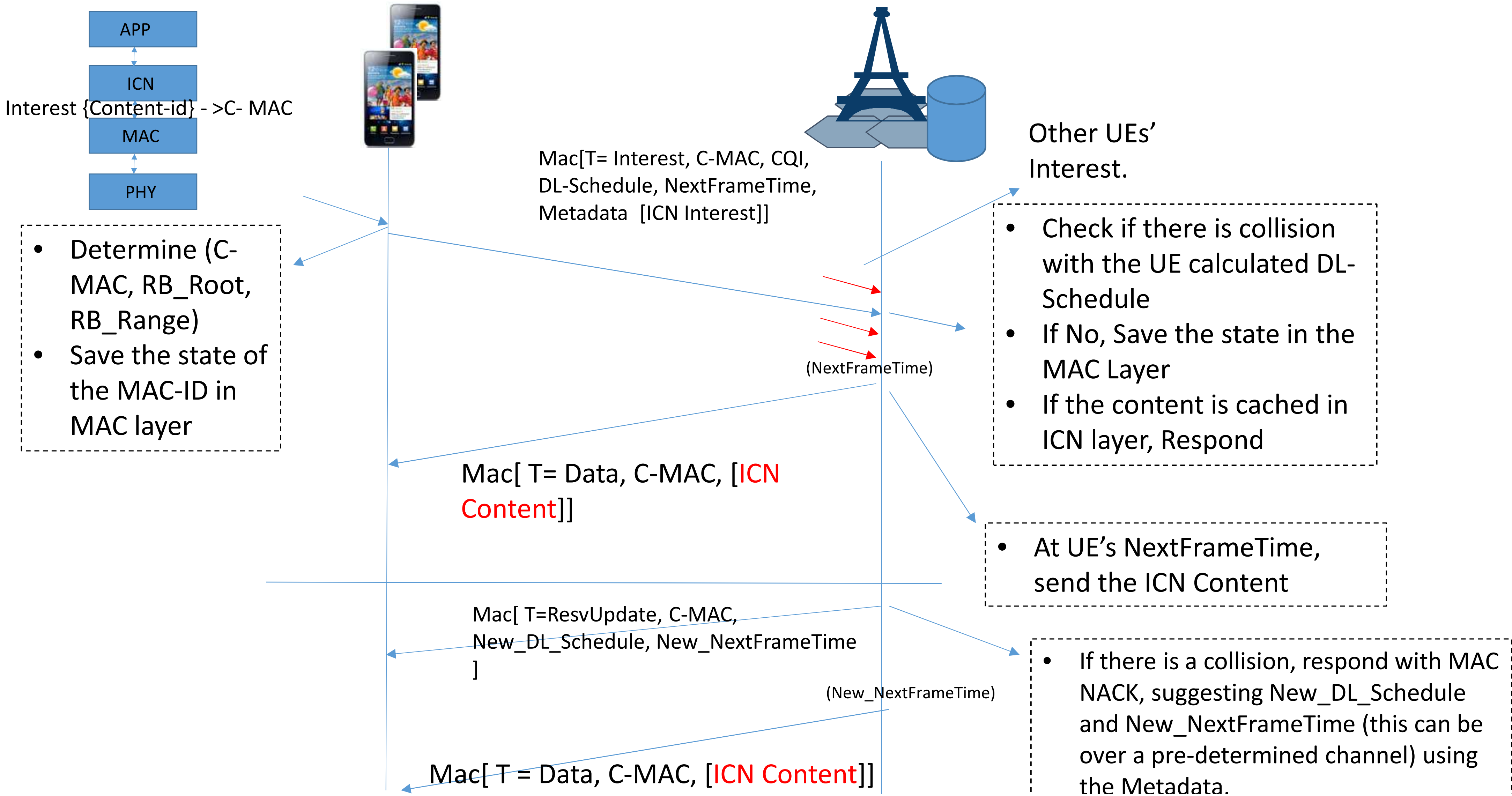
S3: Now determine NextFrameTime, which is a function of the estimated RTT and the delay considering ICN/MAC/PHY layers processing time.

- One approach is to set NextFrameTime considering the considering cached/stored in the eNodeB.
  - NextFrameTime = CurrentFrameTime + Rounded_NextSubFrameTime(Estimated_RTT).
- NextFrameTime is the next LTE sub-frame when the receiver will listen for the response
- NextFrameTime should also consider the case when Interests will be spaced from multiple UEs within a subframe
- NextFrameTime should be invariant to UE request time to ensure consensus on multicast schedule

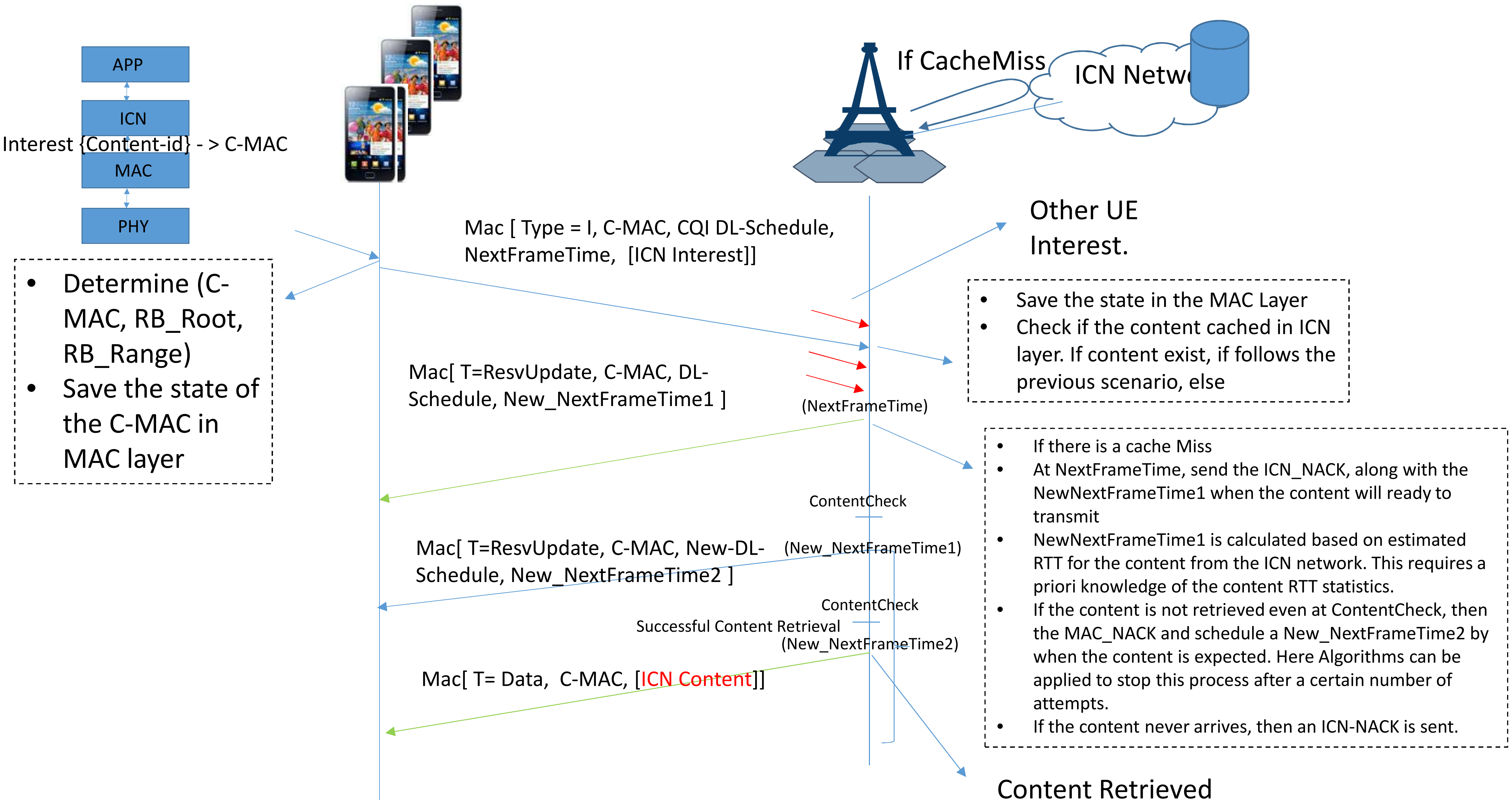
→ Depending on the Channel Capacity and traffic characteristics there can be significant collision on the RB resource choice by the UEs for different content-IDs.

We resolve this at the eNode-B as we discuss this later.

Protocol between UE and eNodeB with Content Cached at eNodeB

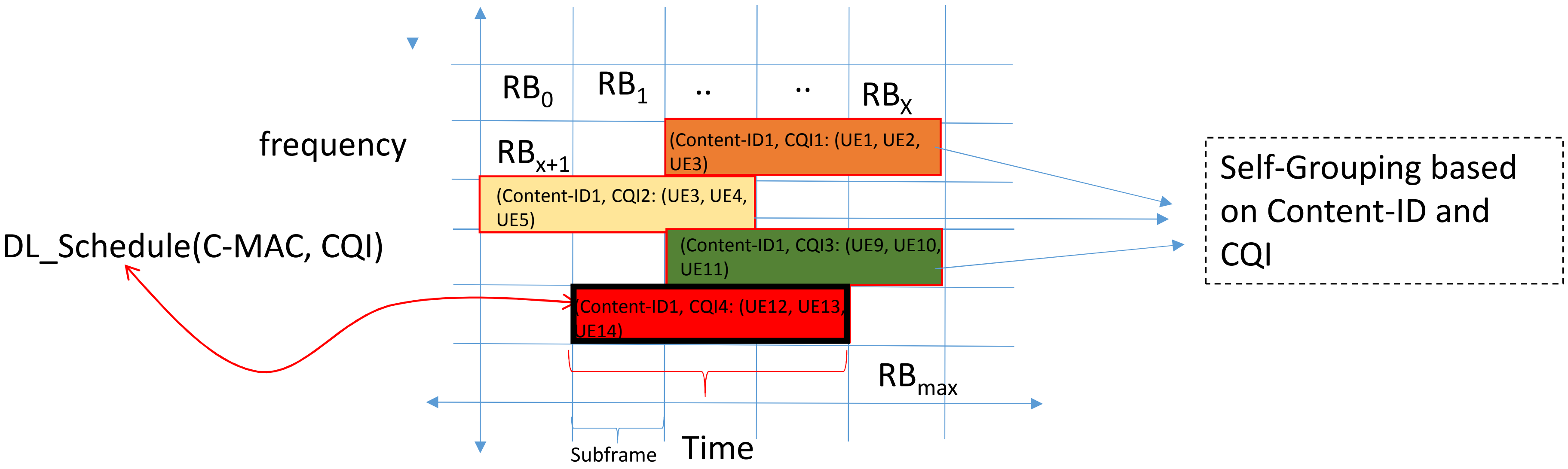


Protocol between UE and eNodeB with Cache Miss

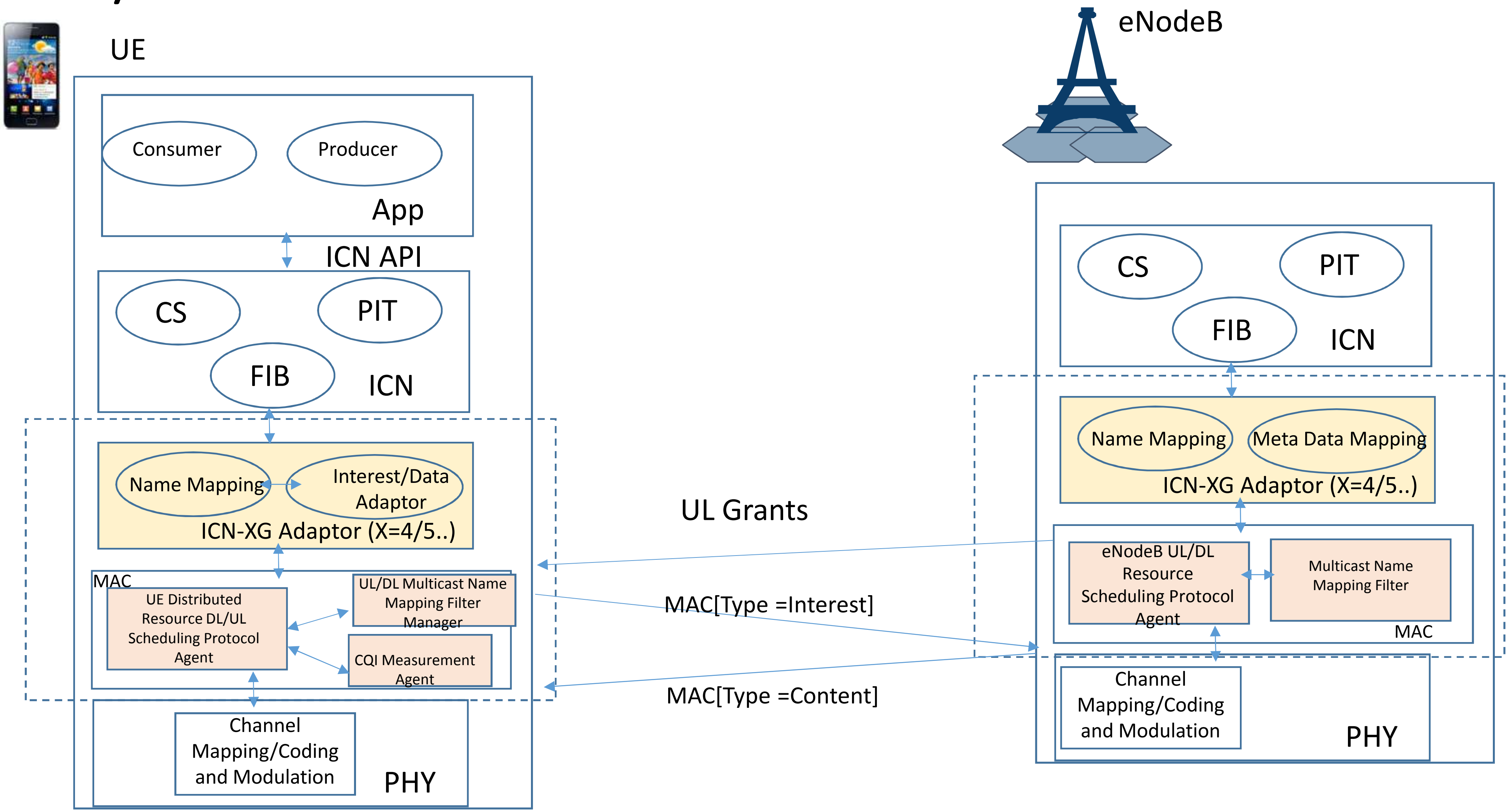


Dynamic Self-Grouping based on CQI

- The mapping from C-MAC to RB can also take the CQI into account to enable self-grouping.
- This is useful as static users will enjoy better CQIs than the mobile ones.
- This is important because different group of UE's could be in different radio conditions.
  - The goal here is to achieve physical separation in the allocation of the physical resource among the group of UEs with different CQIs
  - *DL_Schedule*(C-MAC, CQI)= f(C-MAC, CQI, RB_{MAX})



System Architecture



- Name Mapping Maps the ICN name to C-MAC. Interest/Data Adaptor uses the API offered by the MAC layer to send and receive Interest/data objects along with any ICN metadata for MAC scheduling
- Multicast name filter holds the state of the names (outstanding CMAC) expressed by applications and those that received from the other users.

Persistent DL Scheduling

- The proposal till now computes C-MAC for each new Content-ID.
- This can be optimized by taking advantage of the context in names, e.g. a C-MAC can be same for set of all chunks belonging for entire content, e.g. a movie
  - i.e., /disneymovie/wonderwomen - > C-MAC, hence /disneymovie/wonderwomen/chunk-x - > C-MAC -> {DL RB Set}
- With some enhanced control channel signaling, this C-MAC and mapping to RB-set can be sent once and avoid the control channel cost, unless the CQI of the UE changes.

Handling C-MAC Hash Collisions

Physical isolation of hash ID based on name space

- Several optimizing techniques can be used to map ICN names to unique hash ID to minimize collision
- Names being contextual, hence can use different hash functions or use one or multiple components as an input to the hash function to physically separate the hash space.

Brute force Handling at eNodeB:

- If two different content IDs map to the same Hash ID, the resolution can be handled at eNodeB's ICN layer.
- The ICN layer can inform the MAC layer of the collision. The MAC can then generate two different C-MACs multicast ID, and this mapping can be sent in the NextFrameTime instance determined by the UEs for the two content-IDs.
- The UEs can then re-request the content with the new C-MACS

Opportunities

- **Static and Mobile Broadband**
  - These solutions can serve both static (residential broadband) and mobile users
- **Network Slicing and CloudRAN**
  - NS allows feasibility of new RAT with new CP and UP features
  - ICN based over cellular access along with CloudRAN intelligence could help in handling large scale content multicast delivery applying software defined radio principles.
- **Cross Layer Optimization**
  - Use the rich contextual information in the ICN layer in the MAC layer opens up new opportunities for cellular operators
  - Information-centric MAC allows to introduce new kinds of prioritization and scheduling, e.g. QoS
- **Several Application Scenarios**
  - Live and VoD Content
  - Pushing out software updates
  - AR/VR
  - Etc.

Challenges

- **Shrinking cell size : Femto, PICO, Small Cells**
  - Study these ideas within HetNets Evironments
    - Valuable for very high bandwidth applications (e.g. AR/VR) within home or spanning few homes
- **Incremental Design**
  - Realizing these designs within LTE or 5G-NR frameworks
  - Would radio slicing allow orthogonal MAC/PHY realizations
- **PHY Reliability**
  - Considering Mobility and device fading scenarios
  - New forms of FEC/HARQ/MCS Mechanism

Back UP