5G phase-1

PHY/MAC design concepts of 5G

Raphael Visoz
Outline

- Introduction
  - Background (standardization process, requirements/levers, LTE vs 5G)

- Part I: 5G PHY/MAC Enablers
  - Physical channels, physical reference signals
  - Frame structure/numerology
  - Waveform
  - Massive MIMO
  - Synchronization
  - Beam management

- Part II: 5G Design principles
  - Forward compatibility
  - Lean design
  - Stay in the box
  - Avoid strict timing relations
  - TDD and FDD design
  - Low latency

- Conclusion
Introduction

- The December 2017 deadline for the first set of New Radio (NR) technical specifications, called early drop out/December acceleration, was finally achieved
  - The December acceleration was initially motivated to catch up with proprietary fixed wireless access solutions in mmW
  - The early drop out scope is limited to Non Stand Alone (NSA) NR, i.e., Dual Connectivity (DC) with LTE
  - No major impact on L1/L2 is expected before the phase 1 finalization (or NR Rel. 15) planned for June 2018 and including SA and NSA NR
  - To cope with the workload RAN1 meetings have dramatically expanded in terms of number of delegates (up to 600 RAN1 delegates), number of contributions submitted (up to 2800 contributions) and parallel sessions (2 NR and 3 LTE-A pro) with myriad of anarchic offline sessions…
Introduction

NR phase 1 inherits many concepts and techniques from LTE since the principle of CP-OFDM based waveform and OFDMA multiple access remain unchanged.

NR phase 1 opens the degrees of freedom of the MAC/PHY layer of LTE in order to cater for

- A wide variety of services (eMBB, URLLC, mMTC)
- Higher frequencies (mmW)
- Wider bandwidth (400 MHz, ~1GHz with CA up to 16 CCs)
- Higher number of antennas (Massive MIMO)

The main levers (phase 1) considered to answer the ambitious goals of 5G (initially set by the METIS project)

- More bandwidth, more antennas, more base stations
- Issues: cost, acceptability by the public (EMF exposure)

NR - New Radio
PART I: 5G PHY/MAC ENABLERS
Physical Channels

**PDSCH**
DL shared channel

**PBCH**
Broadcast channel

**PDCCH**
DL control channel

**DL Physical Signals**
Demodulation Ref (DMRS)
Phase Tracking Ref (PT-RS)
Tracking Ref (TRS)
Ch State Inf Ref (CSI-RS)
Primary Synch (PSS)
Secondary Synch (SSS)

**PUSCH**
UL shared Channel

**PUCCH**
UL control channel

**PRACH**
Random access channel

**UL Physical Signals**
Demodulation Ref (DMRS)
Phase Tracking Ref (PT-RS)
Sounding Ref (SRS)

Physical data channels (PDSCH/PUSCH) are CP-OFDM based configured with a given numerology
## Physical channels

- **DL physical channels**

<table>
<thead>
<tr>
<th></th>
<th>PDSCH (5G)</th>
<th>PDSCH (LTE)</th>
<th>PDCCH (5G)</th>
<th>PDCCH (LTE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
<td>Transmit DL Data</td>
<td>Transmit DL Data</td>
<td>L1/L2 Control channel</td>
<td>L1/L2 Control channel</td>
</tr>
<tr>
<td><strong>Waveform</strong></td>
<td>OFDM*</td>
<td>OFDM</td>
<td>OFDM*</td>
<td>OFDM</td>
</tr>
<tr>
<td><strong>Bandwidth</strong></td>
<td>Numerology dependent</td>
<td>1.4/3/5/10/15/20 MHz</td>
<td>Localized in BWP</td>
<td>Spread out in the entire bandwidth</td>
</tr>
<tr>
<td><strong>Reference signals</strong></td>
<td>Only UE specific signals (DMRS)</td>
<td>Cell specific or UE specific (Rel. 10)</td>
<td>UE specific (DMRS)</td>
<td>Cell specific (CRS)</td>
</tr>
<tr>
<td><strong>Modulation</strong></td>
<td>Up to 256QAM</td>
<td>Up to 256 QAM</td>
<td>QPSK</td>
<td>QPSK</td>
</tr>
<tr>
<td><strong>Coding scheme</strong></td>
<td>LDPC</td>
<td>Turbo</td>
<td>Polar</td>
<td>TBCC</td>
</tr>
</tbody>
</table>

* With filtering or time domain windowing
# Physical channels

- **UL physical channels**

<table>
<thead>
<tr>
<th></th>
<th>PUSCH (5G)</th>
<th>PUSCH (LTE)</th>
<th>PUCCH (5G)</th>
<th>PUCCH (LTE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
<td>Transmit UL Data</td>
<td>Transmit UL Data</td>
<td>L1/L2 Control information</td>
<td>L1/L2 Control information</td>
</tr>
<tr>
<td><strong>Waveform</strong></td>
<td>OFDM* or DFT-s-OFDM*</td>
<td>DFT-s-OFDM</td>
<td>Filtered OFDM* or DFT-s-OFDM*</td>
<td>DFT-s-OFDM</td>
</tr>
<tr>
<td><strong>Bandwidth</strong></td>
<td>Depend on numerology</td>
<td>1.4/3/5/10/15/20</td>
<td>Many flexible formats in time/freq.</td>
<td>1 RB in freq. 14 symbols time</td>
</tr>
<tr>
<td><strong>Modulation</strong></td>
<td>Up to 256 QAM (\pi/2)-BPSK</td>
<td>Up to 256 QAM QPSK, (\pi/2)-BPSK</td>
<td>QPSK</td>
<td>QPSK</td>
</tr>
<tr>
<td><strong>Coding scheme</strong></td>
<td>LDPC</td>
<td>Turbo</td>
<td>RM/Polar</td>
<td>RM/TBCC</td>
</tr>
</tbody>
</table>

* With filtering or time domain windowing
Frame structure/numerology

- A **numerology** is defined by a subcarrier spacing and a CP overhead.
- There is a fundamental relationship between the OFDM symbol duration $T$ and the subcarrier spacing $\Delta f$:
  \[ \Delta f \times T = 1 \]
  In NR, $\Delta f = 2^\mu \times 15$ kHz, $\mu=0,1,2,3,4$

- Why higher Sub-Carrier Spacing (SCS) than LTE?
  1. More robust to phase noise and Doppler (mmW)
  2. Better latency since when $\Delta f$ increases, the symbol duration decreases
  3. Wider bandwidth for a given IFFT size
     - LTE default 2048, can reach 20 MHz with 15kHz SCS
     - NR default 4096, can reach 100MHz and 400MHz with 30 and 120 kHz SCS, respectively
Frame structure/numerology

- RAN4 has selected:
  - \(\{15, 30, 60\}\) kHz < 1GHz
  - \(\{15, 30, 60\}\) kHz \(\in [1,6]\)GHz
  - \(\{60,120, 240 \text{ control only}\}\) kHz > 6GHz
  - General assumption: 30 kHz @3.5GHz, 120 kHz @28 GHz

- Normal CP means that the guard time period to prevent ISI is kept proportional to symbol duration \(T\) (~8%)
  - Small SCS means large CP => can cope with large delay spread (MBMS)
  - Large SCS means small CP => can cope only with small delay spread (mmW)

<table>
<thead>
<tr>
<th>SCS</th>
<th>CP duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 kHz</td>
<td>4.69 (\mu)s</td>
</tr>
<tr>
<td>120 kHz</td>
<td>4.69/8=0.59 (\mu)s</td>
</tr>
</tbody>
</table>
Frame structure/numerology

- **Fame and subframe**
  - The 15 kHz numerology is kept as reference with
    - 1 ms sub-frame
    - 10 ms frame

- **Symbol level alignment**
  - In order to allow symbol level TDM between numerologies

![Diagram showing symbol level alignment and TDM]
Resource grid

PRB alignment for FDM between different SCS

<table>
<thead>
<tr>
<th>µ</th>
<th>Min RB</th>
<th>Max RB</th>
<th>SCS</th>
<th>Tx Bw Min (MHz)</th>
<th>Tx Bw max (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>270</td>
<td>15</td>
<td>3.6</td>
<td>48.6</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>273</td>
<td>30</td>
<td>7.2</td>
<td>98.3</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>264</td>
<td>60</td>
<td>14.4</td>
<td>190.1</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>264</td>
<td>120</td>
<td>28.8</td>
<td>380.2</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>TBD</td>
<td>TBD</td>
<td>TBD</td>
<td>TBD</td>
</tr>
</tbody>
</table>

Note: 20 PRB is the SS bandwidth

RB – Resource Block
PRB – Physical RB
FDM – Frequency Division Multiplexing
SCS – SubCarrier Spacing

2018 – Version 1.0
Frame structure numerology (examples)

Typical @3.5GHz

Typical @28GHz
Waveform

- Two approaches that are RAN1 spec. transparent:
  - Per-subcarrier filtering or time domain windowing: Weighted Overlap and Add
  - Sub-band filtering: Filtered OFDM

- The NR wave form is CP-OFDM based which means that it can be received by a legacy CP-OFDM receiver (without disrupting too much the complex orthogonality between carrier)
- DFT spread/SC-FDMA can be configured in the UL by the network as a PAPR reduction technique
Waveform

- Filtering allows a better Spectral Utilization $SU = \frac{\text{Transmission Bandwidth}}{\text{Channel Bandwidth}}$

  - Can fit more PRBs into a channel bandwidth

  - Allow less guard band between different SCS that are FDM
## Massive MIMO

<table>
<thead>
<tr>
<th>Release</th>
<th>Description</th>
</tr>
</thead>
</table>
| Rel. 8  | 4 antenna ports  
          | 1D antenna array  
          | TM3/4/6/5         |
| Rel. 10 | 8 antenna ports  
          | 1D antenna array  
          | TM9              |
| Rel. 11 | COMP         
          | TM10           |
| Rel. 13 | FD-MIMO (2D antenna array)  
          | 16 antenna ports  
          | Beam management for data class B  
          | TM9/10         |
| Rel. 14 | 32 antenna ports |
| NR      | Superset of Rel. 13/14  
          | Beam management for data and control |

COMP - COordinated Multi Point operation
Massive MIMO

- Massive MIMO is the extension of MIMO with a large number of controllable antennas
  - @3.5GHz: typically 128 antenna elements
    - Large number of antennas increases capacity thanks to spatial multiplexing
  - @28GHz: typically 512/1024 antenna elements
    - Large number of antennas (N) allows space focalization $\sim 10\log(N)$ to fight back pathloss
Massive MIMO

- mmW
  - Things can get more complicated with hybrid (digital–analog) beam-forming

1) Beam management: to find the right analogical beam
2) CSI acquisition: simple PMI feedback

N TXRUs where N is not greater than 4
P TX/RX paths where P very large a few hundreds

g-Node B
4 panels
1 TXRU per panel

UE
1 panel w. 4 TXRUs

2018 – Version 1.0
Massive MIMO

- mmW: why only a few TXRUs?
  - Cost/technology issues
    - Compared to cmwave (3.5 GHz), mmw have their antenna spacing of a few millimeters
      - 3.5GHz = half wavelength/antenna spacing 4cm
      - 28GHz = half wavelength/antenna spacing 5mm
    - it is extremely difficult to have more than one TXRU per panel for space issues (for the different phase shifters, connections and adders).
    - The targeted bandwidth can be very large (400MHz) and it calls for a very high sampling rate that makes the Digital to Analog Converter (DAC) very expensive.
  - Interest
    - MU-MIMO is a capacity improving technique, in mmW we are mostly in power limited regime, i.e., splitting the power between users is not a good idea for this regime.
Massive MIMO

- For both PDSCH and PUSCH, the CSI acquisition can be based on
  - Full channel reciprocity:
    - The estimated UL (DL) channel gives the DL (UL) channel from where the precoding is chosen
      - ZF for MU-MIMO (minimize the interference between served users)
      - Eigenvector Based for SU-MIMO (maximize capacity and reduce interference between spatial layers)
    - The receiver only feedbacks RI and CQI (interference situation), however the RI and CQI derivations depend on the selected precoding
      - DL: Base station indicates to the UE the CSI-RS ports pre-coded with the chosen precoder
      - UL: Base station deduces the precoder from precoded SRS ports
Massive MIMO

- Codebook based PMI feedback
  - The UE feedbacks the **Precoder Matrix Indicator** (PMI), **Rank Indicator** (RI) and **Channel Quality Indicator** (CQI)
  
  - The UL codebooks are very simple limited to rank 1-2 with up to 4 antenna ports and inherited from LTE
  
  - The DL codebooks are based on the **W1W2** structure allowing very large antenna array (2D Uniform Linear Antenna array for each polarization) with up to 32 ports
  
  - **W1** select the transmission direction in elevation and azimuth (long term) for both polarization
  - **W2** co-phase the receive polarization to add them coherently
Massive MIMO

1. Beam-forming gain are needed to fight back path-loss @mmW
   - All physical signals must be beam-formed
2. Beams cannot reach all the users due to their directivity

How do we deal with Broadcast signals? Solution Beam sweeping/switching
Synchronization

- NR follows a beam centric approach: All physical channels, reference signals are beam-formed
- For carrier frequency range up to 3 GHz, Max number of beams: 4
- For carrier frequency range from 3 GHz to 6 GHz, Max number of beams: 8
- For carrier frequency range from 6 GHz to 52.6 GHz, Max number of beams: 64
- SS blocks are gathered within 5ms in specific OFDM symbol positions

![Diagram of SS blocks and OFDM symbol number]

SS-Synchronization Signal
TRP-Transmit Receive Point
Beam management

1. Beam sweeping at Tx for TRP and at Rx for UE to align transmit and receive beams: **Beam pair link**

2. UE reads PBCH/RMSI on that beam, RMSI indicates associated PRACH resource (with same receive beam as the transmit one)

3. UE based on **beam correspondence** send the PRACH on the indicated resource

4. RRC connection

5. Refinement/selection/maintenance of the beam based thanks to precoded CSI-RS
PART II: 5G DESIGN PRINCIPLES
Design principles: Forward Compatibility

• Agreement on forward compatibility (first NR RAN1 meeting):
  
  – **5G will follow a two-phase approach, the first phase aims at mid 2018**
    – Phase 1 and later phases of NR should be designed with the following principles to ensure forward compatibility and compatibility of different features:
    – **Strive for**
      – Maximizing the amount of time and freq. resources that can be flexibly utilized or that can be left blanked without causing backward compatibility issues in the future (avoid fixed reference signal except for synchronization if necessary)
        – Blank resources can be used for future use
      – Minimizing transmission of always-on signals
    – **Confining signals and channels for physical layer functionalities (signals, channels, signalling) within a configurable/allocable time/freq. resource**
Design principles: Lean Design

- **Lean design**: minimize “always on” transmission for forward compatibility and network energy efficiency
  - LTE always on signals: Synchronization Signal (SS/5ms periodicity), Cell specific reference signals (CRS), broadcast system Information
  - NR: No CRS, principle of configurability and on demand transmission (in connected mode)
    - Examples:
      - Configurable SS periodicity {5, 10, **20** (default), 40, 80, 160} ms
      - Configurable fine time/frequency tracking reference signal (TRS)
      - On demand Other System Information (OSI)
Design principles: Stay in the Box

- Stay in the box for forward compatibility and narrow band UE capability handling
  - **LTE**: Some control channels in LTE are spread out wideband (PCFICH/PHICH/PDCCH) which makes introducing new transmissions difficult in the control region of LTE
    - NB-IoT avoids LTE control regions
    - e-MTC had to redesign the PDCCH due to its Narrow Band capability
  - **NR**: Introduces the concept of BandWidth Part where the control and data should be contained in frequency within a bandwidth part of a wider CC
Design principles: Avoid Strict Timing Relations

- Avoid strict timing relations for forward compatibility and latency reduction

LTE FDD: Fixed timing relations between
- PDSCH and ACK (n+4)
- UL grant and PUSCH (n+4)

The UE has fixed 3ms -TA processing time budget

These fixed timing relations are detrimental for:
- **Latency** ~8ms HARQ RTT (cannot adapt to better UE capability)
- A sub-frame at time n in UL (DL) cannot be left blank if a transmission occurred at time n-4 in DL (UL)
Design principles: Avoid Strict Timing Relations

- LTE TDD: configuration 2 (4:1)
  - Special sub-frames are bidirectional sub-frame/slot with a downlink part carrying a shortened PDSCH (DwPTS), a Guard Period (GP), and an uplink part for channel sounding or short PRACH (UpPTS denoted U below), e.g., special sub-frame 7 with DwPTS=10, GP=2 and UpPTS=2 symbols.
Design principles: Avoid Strict Timing Relations

- LTE TDD with reference configuration 2 (4:1)
  - Downlink association set \{k0, k1, k2, k3\} dictated by the constraint that an ACK associated to a PDSCH received at subframe n cannot be sent before subframe n+4
  => PDSCH at n-8, n-7, n-6, n-4 are acknowledged at UL subframe n
Design principles: Avoid Strict Timing Relations

- **Definition:**
  - K1: Delay in Time Transmission Interval (slot) between DL data (PDSCH) reception and corresponding ACK transmission on UL, e.g., in LTE $K1 \geq 4$ TTI
  - K2: Delay in TTI between UL grant reception in DL and UL data (PUSCH) transmission, e.g., in LTE $K2 \geq 4$ TTI

- In NR there are no fixed timing relations:
  - K1 and K2 can be dynamically adapted by the network to the UE processing time capability, Timing Advance as well as DL:UL ratio and switching points
  - K1=0 defines a **self contained slot** in NR for TDD, i.e., the PDSCH and its ACK are contained within a bidirectional slot (very important to fit into a Maximum Channel Occupancy Time in unlicensed spectrum)
Design principles: Avoid Strict Timing Relations

Example: TDD@3.5 GHz with 30 kHz SCS and 3 symbol GP with DL-unknown-UL periodicity equal to 2ms (unknown means Guard Period and/or symbols that can be dynamically allocated to UL or DL)

UE processing allowing $K_1 \geq 2$ ($N_1 = 10$ symbols, long PUCCH)

UE processing allowing $K_1 \geq 1$ ($N_1 = 2-3$ symbols, long PUCCH)
Design principles: TDD vs. FDD

- Maximize the commonality between UL and DL as well as FDD and TDD

- TDD UL/DL configuration can be either semi-static (configured per cell/UE semi-statically) or dynamic. In the dynamic case, the UL/DL ratios, number of switching points, can be changed and indicated by Slot Format Indicator (carried by the common group PDCCH)
  - However, dynamic TDD is not seen as practical for macro deployment
  - to avoid UL to DL interference the network has to be synchronized with same TDD configuration per cell
Design principle: Low Latency

- Low latency targets in terms of one way User Plane latency
  - eMBB below or equal to 4ms
  - URLLC below or equal to 0.5/1ms
    “A general URLLC reliability requirement for one transmission of a packet is $1 \times 10^5$ for 32 bytes with a user plane latency of 1ms.” (TR 38.913)

- Low Latency Communication (LLC) levers
  - Reduced processing time at the UE (2-3 symbol targeted)
    - Highly parallelized LDPC codes (main reason given for its selection for data channels)
    - Single antenna port transmission for transmit diversity (precoding cycling, cyclic delay diversity etc…) transparent to the UE
    - Front loaded DMRS, DL control information in the first symbols of a slot/minislot
    - Resource mapping following (i) spatial layer -> (ii) frequency-> (iii) time to allow pipelining decoding per OFDM symbol
**Design Principle: Low Latency**

- **Low Latency Communication (LLC) levers**
  - Both UL and DL, frame structure with larger SCS and non-slot based scheduling (or mini slots of 7/4/2 symbols)
    - Extended CP with 60kHz for macro deployment (similar CP duration as in LTE 4.17us vs. 4.69 us)
  - For UL, scheduled transmission
    - Short PUCCH format and frequent SR transmission opportunities
  - For UL, grant-free transmission related design with K repetitions
    - Related to LTE Semi-Persistent Scheduling (SPS)
    - The TB repetitions can start flexibly during the K transmission occasions within the Periodicity
    - UL grant/DCI can occur during the K repetitions either to serve as an early ACK (FFS) or schedule the retransmission of the same Transport Block (GF2GB)
**Design Principle: Low Latency**

- Low Latency Communication (LLC) levers
  - For DL, pre-emption indication and Code Block Group (CBG) retransmission
    - Note: a Transport Block (L2 SDU) is transmitted into several code blocks, each code block are encoded separately and can be decoded independently.
    - The g-Node B decides to preempt radio resources allocated to some ongoing eMBB transmission
    - The punctured resources are identified based on the Preemption Indication (PI) carried in Group Common PDCCH DCI next slot
    - Resources received at the UE which are also indicated by the PI are flushed (erasure)
    - Only the code-blocks missing are retransmitted: Code Block Group retransmission
**Design principle: Low Latency**

- **URLLC issues**
  
  - **TDD**: even without retransmission, there should be at least 2 DL/UL switching point during one 0.5ms slot for SCS 30 kHz (to ensure 1ms worst case)
    - User plane latency DL data = TTI (gNB processing) + 2 TTI (frame alignment) + 1 TTI (PDSCH over the air) + UE processing (2OS) = 1.1ms
    - User plane latency (grant free) UL data = UE processing (20S) + 2 TTI (frame alignment) + 1 TTI (PUSCH) + TTI (gNB processing) = 1.1ms
  
  - **FDD**: without retransmission mini slot of 7 symbols can achieve the 1ms worst case latency budget for SCS 30kHz
Design principle: Low Latency

- **URLLC issues:**
  - For URLLC one solution is to retransmit/repeat systematically without waiting any ACK/NACK
    - Solve the reliability of ACK/NACK and RTT delay
    - *Highly inefficient* compared to retransmissions only when needed
      - This can be at least partially solved by *early ACK termination*
Design principles: Low Latency

- URLLC issues
  - URLLC TDD configurations conflict with eMBB TDD configurations for macro deployment in terms of spectral efficiency (high number of switching points, high number of Uplink frames)
  - Solutions:
    - Rely on low frequency FDD band (700MHz): supports a mixed of eMBB and URLLC traffic
    - High frequency dedicated band for URLLC with small cell deployment
      - Unlicensed?
Conclusion

- Phase 1 is more oriented towards the increase of available physical dimensions (e.g., antennas, spectrum, g-Node Bs) rather than the increase of spectral efficiency conditional on fixed resources
  - eMBB is the dominant service targeted by phase 1. Apart from low latency, the verticals (mMTC, UR) will be more addressed during phase 2

- Phase 2 will study
  - Non-orthogonal multiple access
    - User are allowed to transmit on the same time-frequency resources and the number of colliding users can exceed the number of receive antennas
      - Rely on advanced receiver architectures
      - Provide capacity (mMTC), latency (URLLC), robustness to imperfect CSI at Tx (eMBB)
  - Unlicensed spectrum for NR
    - Stand Alone unlicensed access has a lot of momentum to have NR compete with high end WiFi services. New threat and opportunity for operators for B2B
  - Integrated access and backhaul
    - Relaying technologies may allow “low cost densification” with wireless backhaul
  - Satellite communications
    - NR for direct communications between satellite and UEs
Conclusion

- Technical specifications can be found under

- Further reading
  - [1] “4G LTE-Advanced Pro and The Road to 5G”, excellent book on LTE evolution towards 5G by Ericsson
  - [2] [5GmmWave_Webinar_IEEE_Nokia_09_20_2017_final.pdf](5GmmWave_Webinar_IEEE_Nokia_09_20_2017_final.pdf) excellent presentation from Nokia
  - [3] [https://www.ericsson.com/research-blog/lte-latency-reductions-preparing-5g/](https://www.ericsson.com/research-blog/lte-latency-reductions-preparing-5g/) interesting high level presentation on URLLC concepts