RAN Slicing: Challenges, Technologies, and Tools

Navid Nikaein
Professor - Eurecom
OpenAirInterface Software Alliance
Mosaic-5G Initiative

Tutorial at CrownCom 2019
11-12 June, 2019, Poznan, Poland.
Provide an overview of 4G, 5G and RAN slicing

Highlight the importance and timeliness of **softwarization, virtualization, and disaggregation** of RAN to enable **multiservice multi-tenant** RAN toward So-RAN architecture

Cover a **well-balanced research** topics and challenges
5G shall enable
Connected, Controlled, and Flexible Network as a service

Digital Society
Value Creation
Consistent experience
Sustainable business model

5G Promises
LT ➔ Driving Forces of 5G
Communication-Oriented

Today’s 4G is designed for a limited number of UCs

- Throughput-optimized
- Fixed
- Rigid

Is 4G enough?
Monolithic BS
Stateful network entities
Transactional communication mode
Certain level of CP and UP separation
Common entity for user mobility and session management

Communication-oriented 4G
Multi-operator RAN (MORAN)
Shared RAN nodes, dedicated spectrum, but separated CN per operator

Multi-operator CN (MOCN)
Shared RAN nodes and spectrum, but separated CN per operator with proprietary services

Gateway CN (GWCN)
shared RAN and part of core networks

Dedicated core (DECOR)
Deploy multiple dedicated CNs (DCNs) within a single operator network
One or multiple MMEs and SGWs/PGWs, each element

Evolved DECOR (eDECOR)
UE assisted DCN selection
Network Node Selection Function (NNSF) at RAN to select directly the proper DCN towards which the NAS signaling needs to be forwarded
Congestion control and load balancing among multiple DCN with shared MME
Mindful about 3GPPP facts and figures

514 Companies from 45 Countries
50,000 delegate days per year
40,000 meeting documents per year
1,200 specifications per Release
10,000 change requests per year

3GPP R8 Facts and Figures
Evolution from 4G to 5G
5G is not just a new radio/spectrum, but also a new architecture and business helper.

5G: A Paradigm Shift
5G Technology Enablers

- Software Defined Networking
- Fog Computing
- Edge Computing
- SDN/NFV Orchestration
- Network Function Virtualization
- Cloudification Virtualization
- Contextual Networking
- Heterogeneous Networking
- Self Organization Networking
- Ultra dense network
- Advanced MIMO
- Carrier Aggregation of discontinuous bands
- Advanced waveforms
- Flexible and high capacity backhaul
- Millimeter Wave
- Single channel full duplexing
- New Spectrum Allocations
- More Flexible Spectrum

(C) Navid Nikaein 2017
Turn physical infrastructure into multiple logical networks, one per service instance: **One-Network, Many-Service**

**NOT** a one-size fits all architecture **NOT** a Dedicated Network
5G novel usage scenarios: eMBB, uRLLC, mMTC

Multi-disciplinary approach with the fusion of computing, communication, information, and IT

Network Slicing
Multi-service multi-tenant network

© Ericsson WP

Service-oriented 5G
3GPP Role Model (3GPPPP 28.801)

Network Slicing evolves the value-chain of telecom industry:

Decoupling of Players, but the reality might be different

Service-oriented 5G
### 3GPP Re-Architects Mobile Network

<table>
<thead>
<tr>
<th>Feature</th>
<th>3G</th>
<th>4G</th>
<th>5G</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Downlink waveform</strong></td>
<td>CDMA</td>
<td>OFDM</td>
<td>OFDM, SCFDMA</td>
</tr>
<tr>
<td><strong>Uplink waveform</strong></td>
<td>CDMA</td>
<td>SCFDMA</td>
<td>OFDMA, SCFDMA</td>
</tr>
<tr>
<td><strong>Channel coding</strong></td>
<td>Turbo</td>
<td>Turbo</td>
<td>LDPC (data) / Polar (L1 contr.)</td>
</tr>
<tr>
<td><strong>Beamforming</strong></td>
<td>No</td>
<td>Only data</td>
<td>Full support</td>
</tr>
<tr>
<td><strong>Spectrum</strong></td>
<td>0.8 – 2.1 GHz</td>
<td>0.4 – 6 GHz</td>
<td>0.4 – 90 GHz</td>
</tr>
<tr>
<td><strong>Bandwidth</strong></td>
<td>5 MHz</td>
<td>1.4 – 20 MHz</td>
<td>Up to 100 MHz (400MHz for ≥6GHz)</td>
</tr>
<tr>
<td><strong>Network slicing</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>QoS</strong></td>
<td>Bearer based</td>
<td>Bearer based</td>
<td>Flow based</td>
</tr>
<tr>
<td><strong>Small packet support</strong></td>
<td>No</td>
<td>No</td>
<td>Connectionless</td>
</tr>
<tr>
<td><strong>In-built cloud support</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
(1) Ultra-flexible radio-access configurations
   Higher bandwidth
   Higher spectral efficiency (bits/s/Hz/m2)
   Bandwidth parts: tailor bandwidth to UE class (like eMTC narrowbands/widebands)
   Network slicing: new abstractions for service classification down to L1

(2) Compatibility with 4G/5G cores (NSA & SA mode)
   5G dual-connectivity (non-standalone operation)
   Interconnection of evolved 4G eNodeB (ng-eNB) with 5G core

(3) Service-oriented 5G core with cloud-native architecture
3GPP Releases
5G Architecture
Overall 5G Architecture
With 4G Core

Overall 5G Architecture
3 Tier RAN Node

CU0 → DU[0-n] → RRU[0-m]

Functions Split

CP - UP split

Service-Oriented CN

service catalog and discovery

Slice selection function

CP - UP split

3GPP 5G RAN and CN
3GPP Re-Architects Mobile Network

**Slicing Functions**

- NSSF
- NEF
- NRF
- PCF
- UDM
- AF
- AUSF
- AMF
- SMF
- DU
- CU
- UPF
- DN

**Control Plane**

**User Plane**

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMF</td>
<td>Access &amp; Mobility Management Function</td>
</tr>
<tr>
<td>AUSF</td>
<td>Authentication Server Function</td>
</tr>
<tr>
<td>NRF</td>
<td>Network Repository Function</td>
</tr>
<tr>
<td>UDM</td>
<td>Unified Data Management</td>
</tr>
<tr>
<td>NSSF</td>
<td>Network slice selection function</td>
</tr>
<tr>
<td>SMF</td>
<td>Session Management Function</td>
</tr>
<tr>
<td>UPF</td>
<td>User Plane Function</td>
</tr>
<tr>
<td>AF</td>
<td>Application Function</td>
</tr>
<tr>
<td>PCF</td>
<td>Policy Control Function</td>
</tr>
<tr>
<td>NEF</td>
<td>Network Exposure Function</td>
</tr>
</tbody>
</table>

3GPP 5G RAN and CN
Functional Split: RAN & CN
Newly introduced SDAP layer
- Reflective QoS
- Explicit Configuration

4G LTE: 1:1 mapping of EPS bearer to DRB
5G NR: One or more QoS flows may be mapped onto one DRB.
Network Slicing
Network Slicing

- Efficient and adaptive use of radio resources
- No functional isolation

- Functional isolation
- Inefficient use of radio resources

RAN Sharing
(e.g. [NVS – IEEE/ACM TON 2012])

Full Isolation
(e.g. [FLARE – JIP 2017])

Sharing vs. Isolation

© M. Marina
Network Slicing Concept
Multi-service multi-tenant network
1) Softwarization (SDN)
2) Virtualization (NFV)
3) Disaggregation (Cloud)
4) Customization (MEC)
Multi-service multi-tenant network

1) Softwarization (SDN)
2) Virtualization (NFV)
3) Disaggregation (Cloud)
4) Customization (MEC)
Evolution of RAN Slicing

- **RAN**
  - Resource
  - Processing
  - State

- **Cloud computing**
  - RAN sharing (e.g., MOCN)

- **SDN**
  - Multi-tenancy

- **NFV**
  - Software-defined RAN

- **RAN virtualization & disaggregation**

- **RAN part is also evolved from D-RAN to C-RAN**
3GPP re-architects mobile networks

3GPP Network Slicing
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 (1), URLCC (2), MloT(3)

NSSAI is a collection of max 8 S-NSSAI
UE sends NSSAI – based on which related slice(s) are selected
Dedicated or Shared Functions

- Safety/autonomous driving service
- URLLC (Ultra Reliable Low Latency)
- Infotainment/video streaming
- eMBB (Mobile Broadband)
- Maintenance/statistics
- mIoT, low throughput

URLLC Slice
eMBB Slice
Default Slice
mIoT Slice
Dedicated or Shared Resources?
Dedicated or Shared Resources?
Multiplexing Gain

Dedicated or Shared Resources

Benefit: Efficient use of radio resources
RAN Slicing Models
Network slicing definitions
- 3GPP, NGMN, IETF, GSMA

Properties
- Virtual network space
- Customization
- Isolation and sharing
- Programmability

Current State of network slicing
- CN slicing → 3GPP solution
- RAN slicing → Under study

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Deploy RAN</th>
<th>Radio resource</th>
<th>CP function</th>
<th>UP function</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLARE (2017)</td>
<td>D-RAN</td>
<td>Dedicated spectrum allocation</td>
<td>Dedicated</td>
<td>Dedicated</td>
</tr>
<tr>
<td>Rost et al. (2017)</td>
<td>D-RAN</td>
<td>Physical resource sharing</td>
<td>Split into cell and user-specific</td>
<td>Dedicated till real-time RLC</td>
</tr>
<tr>
<td>Ksentini &amp; Nikaein (2017)</td>
<td>D-RAN</td>
<td>Flexible resource sharing</td>
<td>Dedicated</td>
<td>Shared</td>
</tr>
<tr>
<td>ORION (2017)</td>
<td>D-RAN</td>
<td>Virtualized resource sharing</td>
<td>Split into cell and user-specific</td>
<td>Dedicated till PHY</td>
</tr>
<tr>
<td>RAN runtime &amp; C-RAN (2018)</td>
<td>D-RAN</td>
<td>Flexible resource customization &amp; multiplexing</td>
<td>Split into cell and user-specific</td>
<td>Different levels of isolation &amp; sharing</td>
</tr>
</tbody>
</table>
RAN slicing system

(1) Isolate slice-specific control logics while keeping common CP/UP functions

(2) Share radio resources in virtualized or physical form
Components

(1) **Slice context manager** performs lifecycle management of each slice (SLA, active UEs, admission control)

(2) **Virtualization manager**
   - provides a generic form of abstraction for virtualizing radio resources and data plane state
   - presents a virtual/isolated view to each slice virtual control plane

(3) **Radio resource manager** allocates physical resources among slices

(4) **UE association manager** handles slice discovery by UEs and maps UEs to slices
Virtual Control Plane

(1) Interacts with the underlying infrastructure via the virtualization Manager of the Hypervisor - translated into control-data APIs

(2) Operates over vRIB, the locally maintained state of virtual radio resources and data plane - Slice network view and state
ORION RAN Slicing System

Slice 2 controller
Slice 2 controller
QoS optimization
Slice 2 controller

Slice 1 controller
Load Balancing
Slice 1 controller
Load Balancing

Orion Hypervisor
Orion Hypervisor
Orion Hypervisor
Orion Hypervisor
RAN Slicing Execution Env.

(1) Run multiple virtualized RAN module instances with different levels of isolation and sharing.

(2) Pipeline RAN functions to either via multiplexed or customized CP/UP functions.

(3) Share radio resources in virtualized or physical form.
(1) Slice data: Slice context and RAN module context
(2) Context manager: Manage slice data and perform CRUD operation
(3) Slice manager: slice life-cycle, program forwarding engine, conflict resolution
(4) Virtualization manager: resource abstraction, partitioning, and accommodation
(5) Forwarding engine: establish slice-specific UP path
RAN Slicing Model

**Resources** Radio spectrum resources, e.g., carriers, resource blocks, bandwidth

**Processing** Functional block for CP/UP operations, delimited through functional block and described with capabilities

**State** Status of BS CP/UP processing and associated configuration
vRAN

2-level Abstractions

virtual BS sub-network
slicing/services

2nd abstraction

logical BS network
technology

1st abstraction

physical RAN entity
infrastructure

RAN Runtime
vRAN Subnets enabled by 2-level Abstractions

(1) merge
from RU
State S₁, S₂, S₃
Processing C₁, C₂, C₃
Resources R₂
FlexVRAN

(2) create network topology
IBS from RU, DU, CU

(3) split
Res.ProcState
lBS₁
vBS₁,₁ vBS₁,₂ vBS₁,₃
lBS₂
vBS₂,₁ vBS₂,₂
lBS₃
vBS₃,₂
lBS₄
vBS₄,₁
lBS₅
vBS₅,₁
dlBSₙ
vBSₙ,₁
vBSₙ,₃

(4) embed subnetwork
common slice 1 on lBS₁, lBS₂ and lBSₙ → vRAN

5G RU
C₁ = PHY
5G DU
C₂ = MAC, RLC
5G CU
C₃ = PDCP, RRC
F₁

RAN Runtime
Virtual-RAN Example

1. Heterogeneous, disaggregated RAN

2. Homogeneous view for infrastructure owner

3. Embedded slices, infrastructure owner view

4. Embedded slices, slice owner view

RAN Runtime
Slice Enforcement
Spectrum Isolation or sharing
Spectrum scarcity
Spectrum heterogeneity
Spectrum efficiency

5G integrates heterogeneous spectrums

New applications, particular bands/numerology
Process policies defined by various stakeholders
Customized programmable control logics per applications

"enb": [
  {
    "freq_min": val,
    "freq_max": val,
    "max_tx_power": val,
    "fdd_spacing": val,
    "band": val
  }
]
Network Graph

1) Policy graph: sharing agreement/policy between MNOs - LSA and CoPrimary

2) Topology (T) graph: geographic relationship of operators’ cells regarding spectrum sharing

3) Spectrum sharing subgraphs (P+T): match offering cells (e.g. B1:O) and requesting cells (e.g. A1:R)
1) Flexible Numerology
2) Bandwidth Parts
3) Abstraction

<table>
<thead>
<tr>
<th>Requested resources</th>
<th>Abstraction types (Resource granularity)</th>
<th>DL allocation type</th>
<th>UL allocation type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource Block</td>
<td>vRBG Type 0 (Non-contiguous)</td>
<td>Type 0, Type 1, Type 2 distributed</td>
<td>Type 1</td>
</tr>
<tr>
<td></td>
<td>vRBG Type 1 (Contiguous)</td>
<td>Type 0, Type 2 localized</td>
<td>Type 0</td>
</tr>
<tr>
<td></td>
<td>vRBG Type 2 (Fixed position)</td>
<td>Type 2 localized</td>
<td>Type 0</td>
</tr>
<tr>
<td>Capacity</td>
<td>vTBS Type 0 (Min RBG granularity)</td>
<td>All Types</td>
<td>All Types</td>
</tr>
</tbody>
</table>

Resource Slicing
4 Steps to radio resources abstraction:

1. Aggregation
2. Partitioning
3. Virtualization
4. Polling
5. Slice resource allocation
6. Slice Scheduling & Accommodation
7. Multiplexing/preemption
Inter-Slice Resource Partitioning and Polling

3 Slices
1. Slice-specific scheduling, SRM
2. Dynamic Resource partitioning, RM
   - Enforce policy over time

Resource Slicing
Slice programmability: Service differentiation via RRM policy enforcement
Resource Slicing
Multiplexing Gain

Resource Slicing

- High traffic arrival rate
- Low traffic arrival rate

Multiplexed slice resource for further utilization

Unused for multiplex
Function Customization in Monolithic BS

Process Slicing
Function Customization in Disaggregated BS

Process Slicing
1) Graph partitioning to embed a slice state

2) Subgraphs per a slice
   - Shared BS-common
   - Isolated slice-specific

3) State sharing among slices through graph merging
Conclusion
Fusion of Computing, Information and Cellular technologies

(a) 5G and beyond is not only New Radio and verticals, it is also an evolution in General-Purpose computing for wireless networks

(b) More and more software technologies (NFV, SDN, MEC) and Data (mining, analytics) jointly with radio signal processing

Conclusion
3GPP 5G Network started from Rel. 15 phase 1

Flow-level QoS
Heterogeneous and disaggregated spectrum and RAT
Multi-level Network Slicing
Cloud Native Architecture
Open Interfaces and Edge computing

Conclusion
RAN slicing is an on-going research with several challenges: Isolation, Sharing, Customization.

Satisfy requirements of slice owner and operator/infra. provider
OpenSource tools: OAI and M5G Ecosystem

http://mosaic-5g.io/

https://www.openairinterface.org/

Conclusion
Data-driven network control, orchestration, and management

Reason-Predict-Control is a generic framework

Performance is limited by the available computing resources
5G System:
TS23.501 - System Architecture for the 5G System
TS23.502 - Procedures for 5G System
TS29.500 - 5G System, Technical Realization of Service Based Architecture
TS29.501 - 5G System, Principles and Guidelines for Services Definition

5GC components
AMF: - TS29.518 - Access and Mobility Management Services
NRF: TS29.510 - Network Function Repository Services
SMF: TS29.502 - Session Management Services, TS29.508 - Session Management Event Exposure Service
UDM: TS29.503 - Unified Data Management Services

Others:
TS 24.501: Non-Access-Stratum (NAS) protocol for 5G System (5GS)
TS 38.413: NG-RAN; NG Application Protocol (NGAP)
RAN:
3GPP TS 38.401: "NG-RAN; Architecture description".
3GPP TS 37.340: "NR; Multi-connectivity; Overall description; Stage-2".
3GPP TS 38.321: "NR; Medium Access Control (MAC) protocol specification".
3GPP TS 38.322: "NR; Radio Link Control (RLC) protocol specification".
3GPP TS 38.323: "NR; Packet Data Convergence Protocol (PDCP) specification".
3GPP TS 37.324: "NR; Service Data Protocol (SDAP) specification".
3GPP TS 38.331: "NR; Radio Resource Control (RRC); Protocol specification".
3GPP TS 38.133: "NR; Requirements for support of radio resource management".

UE:
3GPP TS 38.304: "NR; User Equipment (UE) procedures in idle mode".
3GPP TS 38.306: "NR; User Equipment (UE) radio access capabilities".

5G RAN Specs.
Personal Info:
Email: navid.nikaein@eurecom.fr
Website: http://www.eurecom.fr/~nikaeinn/
Linkedin: https://www.linkedin.com/in/navidnikaein
Tel: +33.(0)4.93.00.82.11

Mosaic-5G.io :
Mail : contact@mosaic-5g.io
Website : http://mosaic-5g.io
Linkedin: https://www.linkedin.com/in/in/mosaic-5g
Twitter: @mosaic5g

Contact Information

(c) Navid Nikaein 2017