OpenRAN, ORAN Architecture and Use Cases

EE 290-4: Advanced Topics in EE – Wireless Networks
UC Berkeley
04/30/2020

Ravi Ravindran/Kevin Tang
Sterlite Technologies

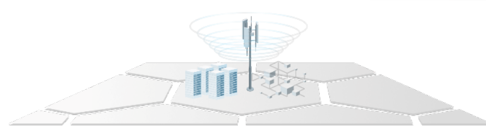

30 April 2020

STL Innovative Solutions Across Layers

AI Driven
Engagement

- AI-powered business
- Intelligence solution
- DevOps-based, Analytical Intelligence-driven, Web-scale-enabled Network Software solution.

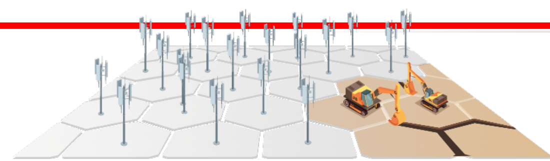
Affordable & Edge
Virtualization



- pRAN is Programmable, Open, Disaggregated Solution
- Enabling programmability at both hardware and software levels

OCP, ONF, TIP, LF, ORAN innovation labs

Hyper Scale & Fast
Deployment

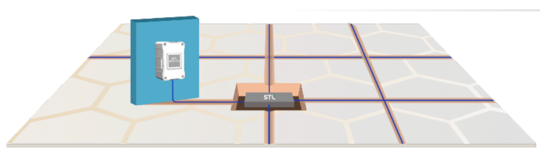


- Ultra fast robotics and AI-based fibre roll outs solution

Algorithmic/ End to End
Design

- Accelerated network
- design solution

Hyper Scale Optical
Connectivity



- Bend insensitive Fibre.
- World's Slimmest Cable.
- Power over Ethernet Innovation

Overview

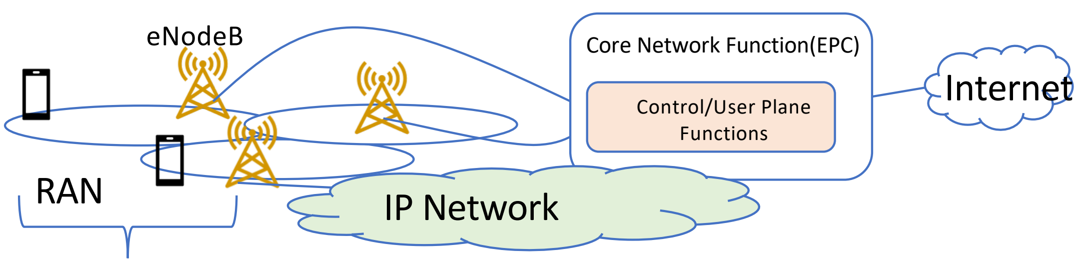
- Partnered with 8 of 10 Top Telcos
- Working with Top 2 Cloud Co.
- Operated in over 100 countries
- $1.5 billion Order Book
- $737 million in revenue
- 43% revenue from international customers

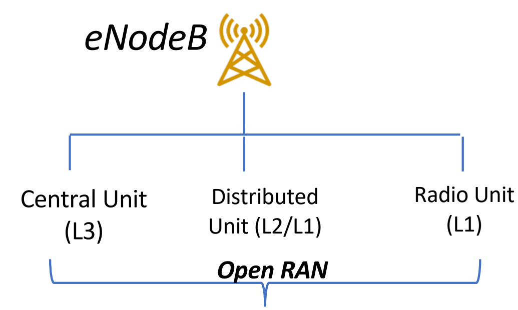
Agenda

- Open RAN Introduction
- ORAN
    - Architecture
    - Near-Realtime RIC
    - Non-Realtime RIC
    - RIC Deployment
- Use case scenarios
    - Overview
    - Traffic Steering
    - Dynamic Spectrum Sharing
- Conclusions

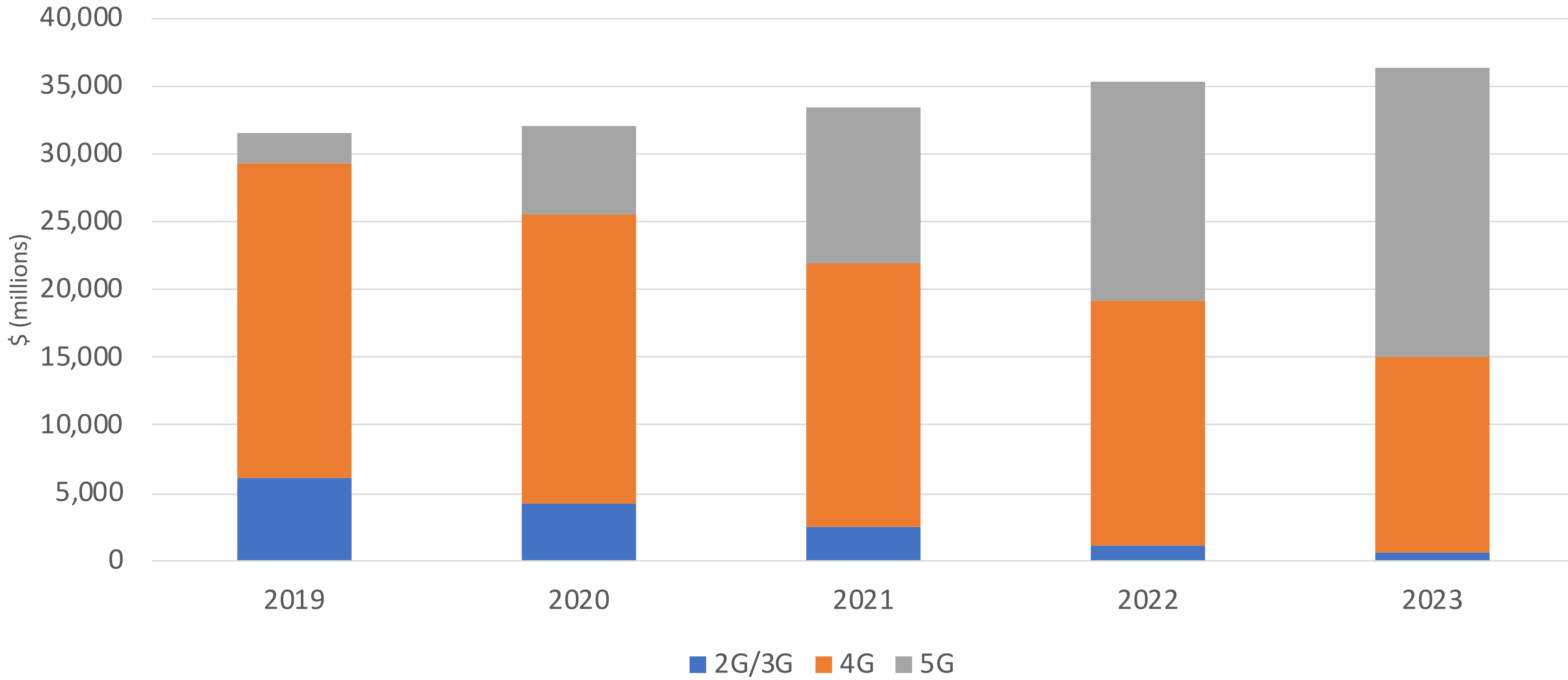
OpenRAN Big Picture

- LTE/5G Network has two main part, the RAN and the Core Network
- RAN is the most expensive segment of a Mobile Operator – next slide
  - Of this owning the spectrum is the biggest slice
- Similar to the well known OSI model, 3GPP based RAN has its own L1/L2/L3 stack
  - For 5g L3 – SDAP/PDCP, L2 – RLC/MAC, L1 – PHY layer
  - All IP traffic between client-serves is tunneled through this stack
- This stack is cross layer optimized to achieve high reliability, mobility, spectrum utilization, throughput, low latency considering the varying channel conditions of the user.
  - Vendors have tried to differentiate using monolithic solutions, hence the cost of this stack.
  - Though 3GPP has standardized the interfaces between these NFs, many procedures such as scheduling, handover algorithms, interference mitigation are closed and proprietary.
- With NFV and SDN, operators want to decouple this stack, hoping to achieve similar performance at a lower cost.
  - This is where OpenRAN becomes increasingly relevant
  - With RAN virtualization the hope is to enable these RAN network function over GPP, and highly available accelerators such as GPU and FPGAs.





RAN Equipment Market Forecast



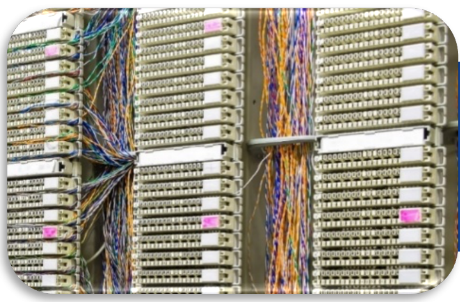
Source: Ovum, 1Q 2019

Lightreading & Sterlite's Webinar:
https://www.lightreading.com/webinar.asp?webinar_id=1488

→ Traditionally monopolized by few vendors such as Nokia, Ericsson and Huawei.

→ Operators are seeking to bring down the RAN cost considering the increasing demand on the services end and decreasing ARPU

Radical shift to Virtualized Radio Access Networks



Open/Virtualized/Programmable RAN

TRADITIONAL

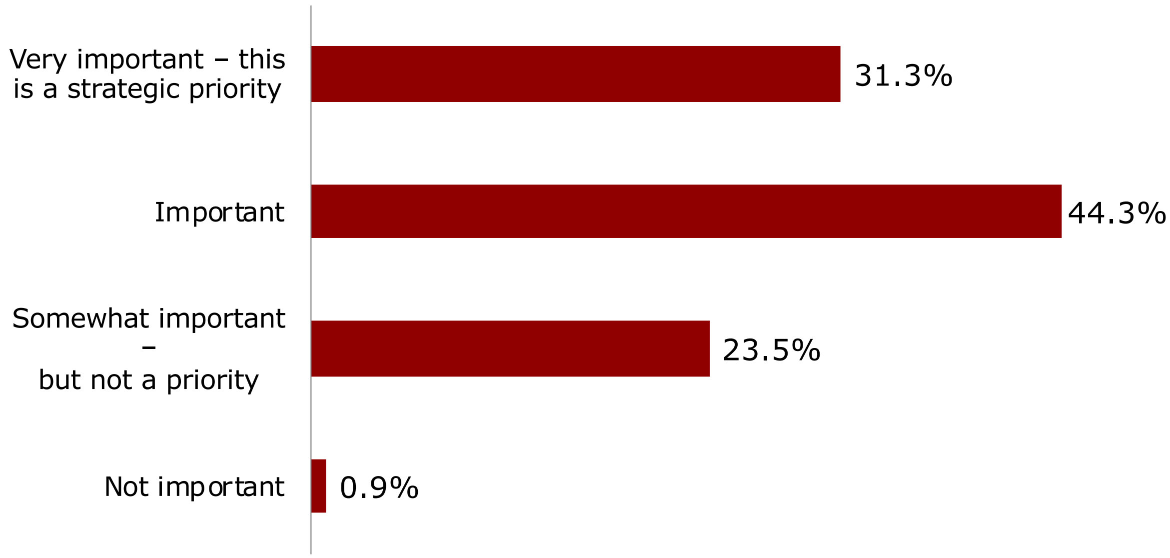
- Closed interfaces
- Vendor specific hardware
- Monolithic and proprietary
- Localized control and data plane
- Expensive

NON-TRADITIONAL

- Standardized open interfaces
- Programmable white boxes
- Cloud native, disaggregated
- Centralized, programmable control plane
- Frugal

Open RAN is Strategic

How important will Open RAN be to your company's network over the next three years? (n=115)



Source: Heavy Reading's Open RAN operator Survey (n=117)

- **A majority of respondents see Open RAN as important or strategic to their company over the next three years.** Optimistic, but not irrationally so, is one way to summarize this response.
- **U.S. respondents were significantly more likely than the rest of the world (RoW) to identify Open RAN as "a strategic priority," (39% vs 25%).** This pattern of the U.S. being more bullish than RoW is repeated several times in the survey.

→ Recently Rakuten a Japanese vendor had a commercial LTE-A launch over fully virtualized implementation, with no major vendors as part of the RAN : - https://www.fiercewireless.com/operators/rakuten-mobile-launches-its-greenfield-network-japan

→ DISH in US is going to be the fourth major operator following the same path, with commercial launch expected in 2023

ORAN Architecture

ORAN Architecture

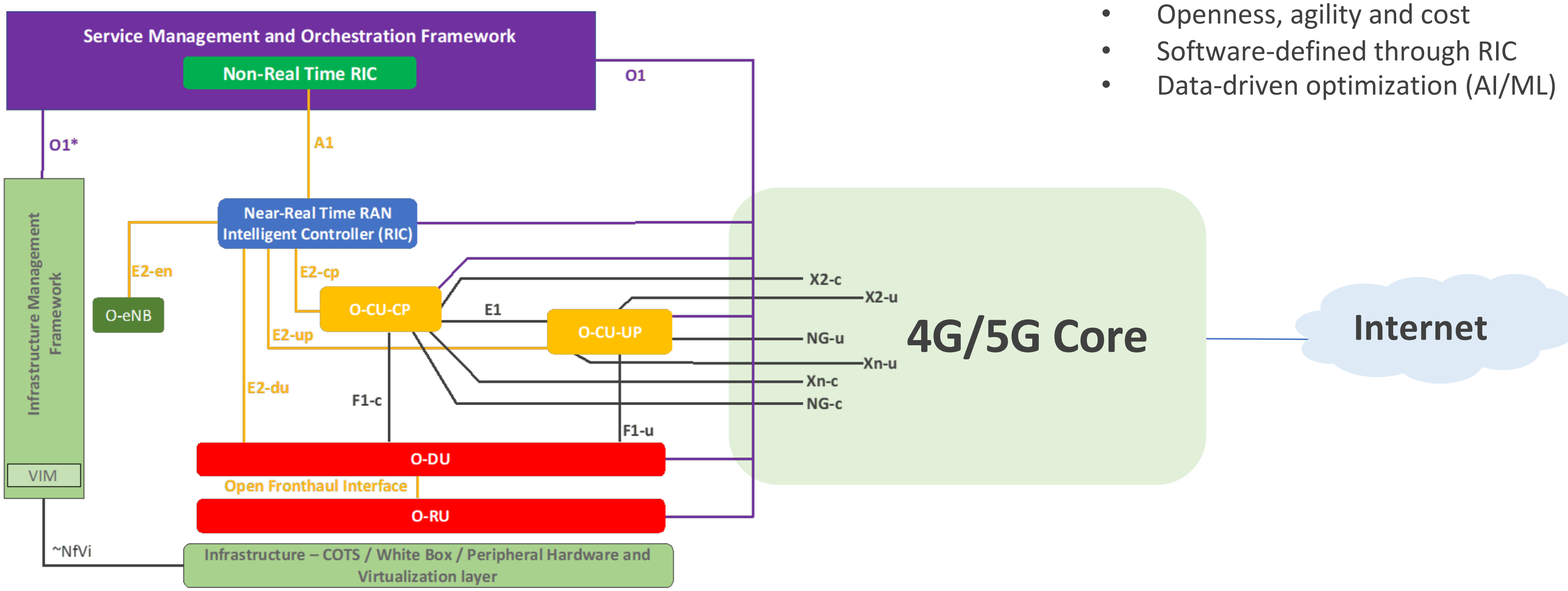


Fig. Source ORAN WG1 - Use case and Architecture

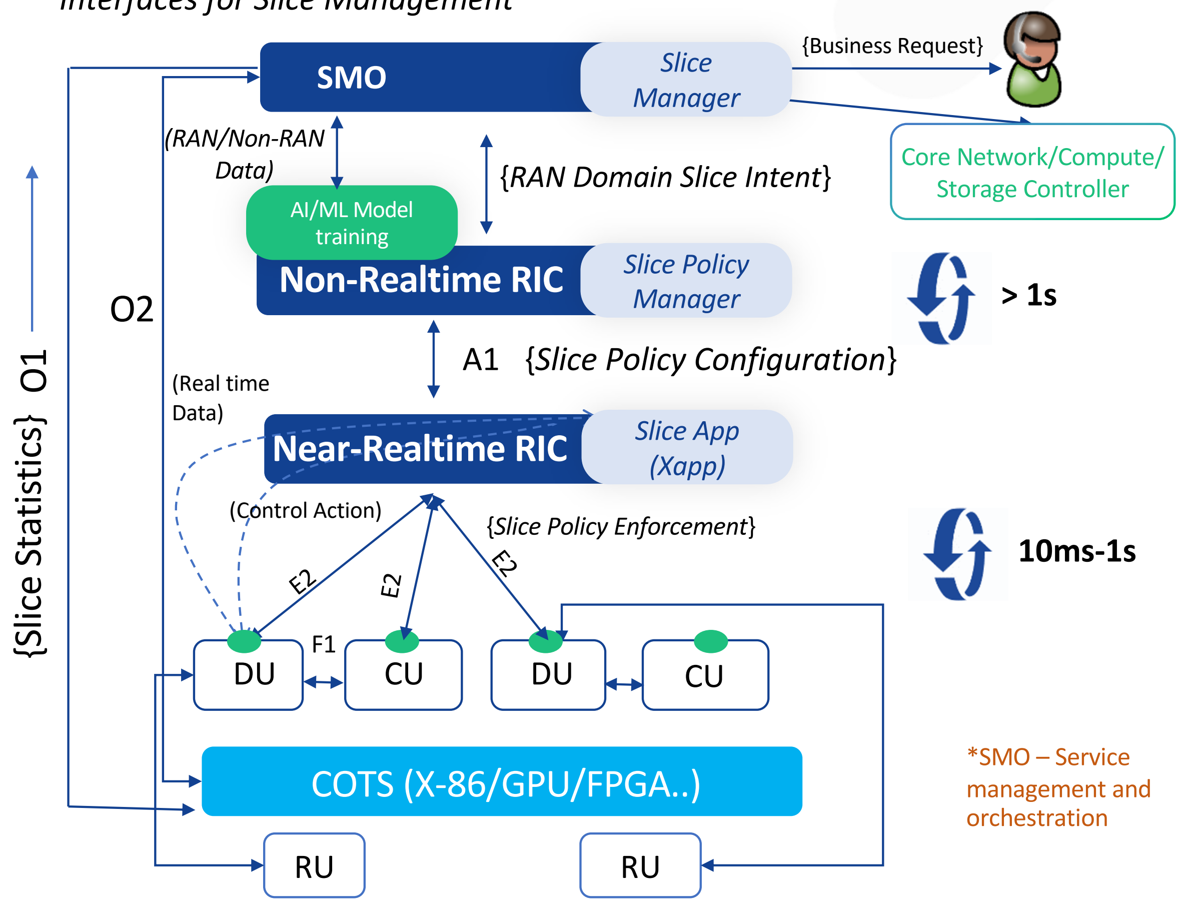
ORAN Architecture :

- RAN focused with 3GPP compatibility
- Openness, agility and cost
- Software-defined through RIC
- Data-driven optimization (AI/ML)

O-RAN Interfaces

- **O2**
  - NFVI related LCM management of virtual E2 nodes or PNFs and RIC components.
- **O1**
  - FCAPS of E2 Nodes, Non-RT and Near-RT RIC
  - Data subscription interface from E2 nodes (DU/CU)
- **A1**
  - Application specific policy, trigger and data management
  - AI/ML model management
  - Application specific enrichment data to near-RT RIC
- **E2**
  - Policy to control mapping
  - Control management
  - E2 node data subscription to X-APPs

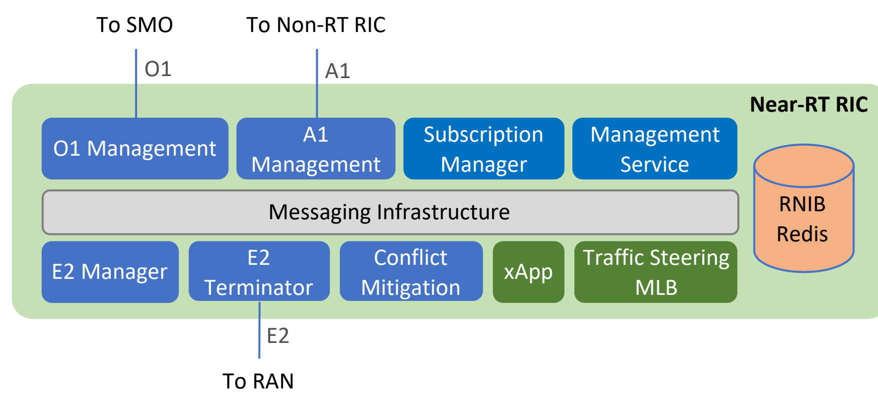
Life cycle showing the use of these three Interfaces for Slice Management



*SMO – Service management and orchestration

Near-Realtime Architecture

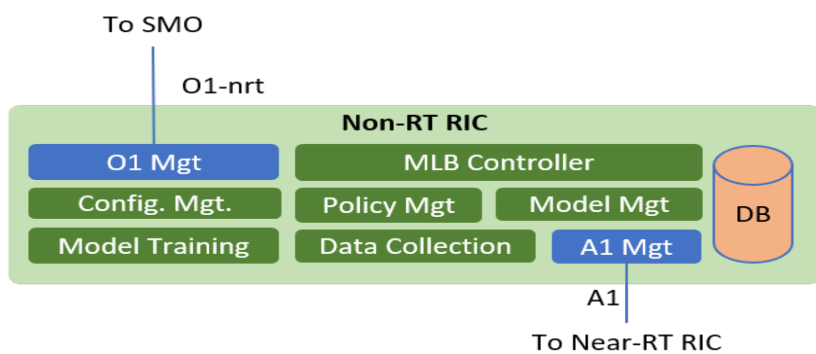
- **Design Considerations:**
  - Operation close to the CU/DU, several deployment choices.
  - Low Latency control Loops (~ms for most stringent applications)
  - Latency sensitive Interface to CU/DU and RU for data subscription and programmability
  - Heterogenous apps, but could also conflict
  - Scaling to hundreds of PMs and events from a single E2 node, in aggregate million/s of events
  - Data Classification and Management (freshness, persistency, caching etc.)
  - Topology Discovery and Management in the RNIB mapping virtual view with physical view
  - RNIB read/write performance for most latency stringent application



High level RIC View

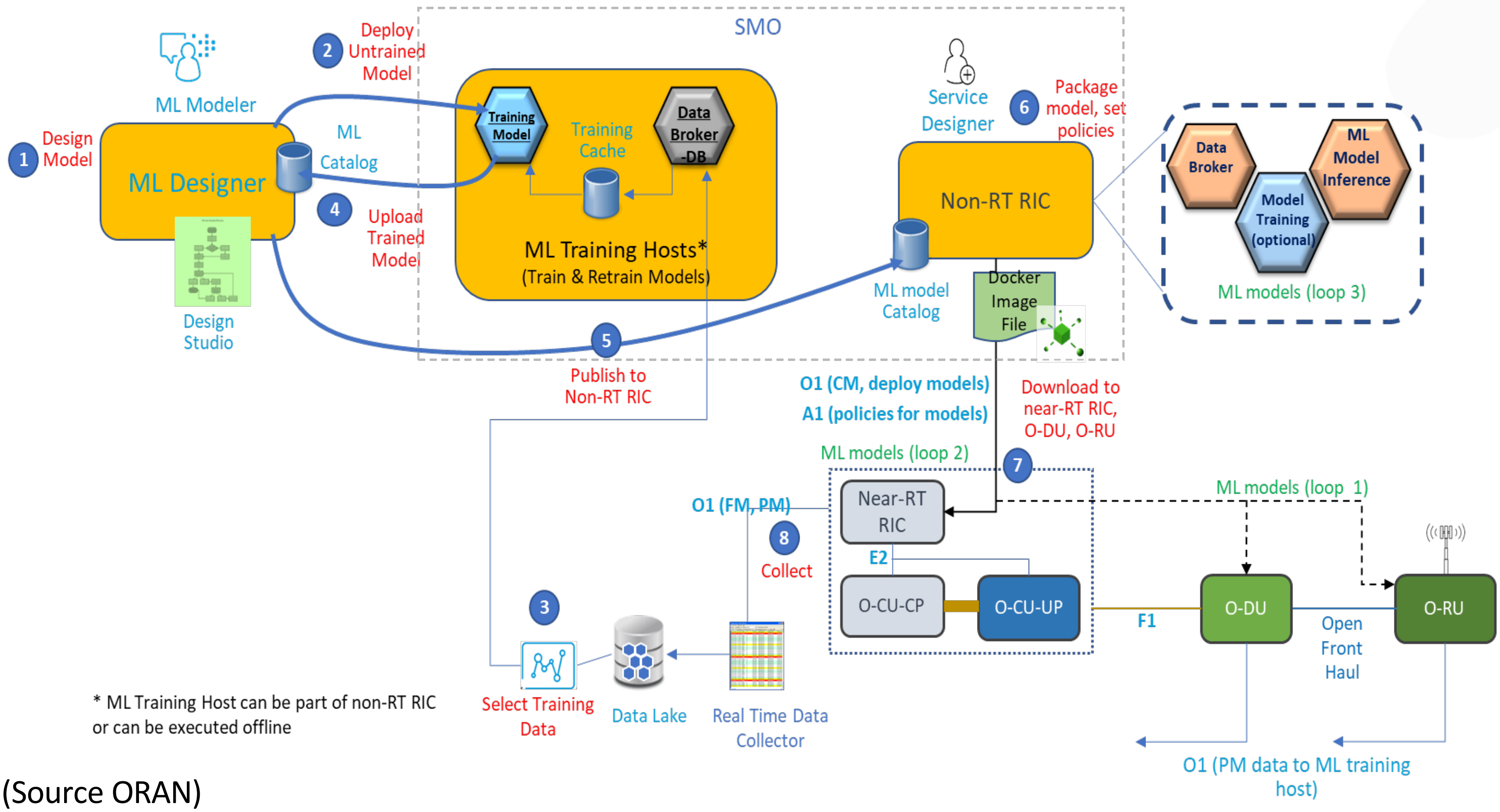
Non-Realtime RIC Architecture

- **Design Considerations:**
  - To be at a higher in the Central cloud
  - R-APPs sets the intent and policies in the near-RT RIC
  - Interfaces with external apps for enrichment data required for Xapps in nRT-RIC
  - Manages many nRT-RICs, hence spans many COs. Should scale to manage PMs and events from several thousands of E2 nodes
  - nonRT-RIC has counterpart application here called R-apps
  - Data management is crucial both for generating policies to nRT RIC and usage for training purposes
  - Interfaces with automated AI/ML platforms for generating trained models for different RAN regions
  - R-Apps to subscribe to PMs over O1



High Level non-RT RIC view

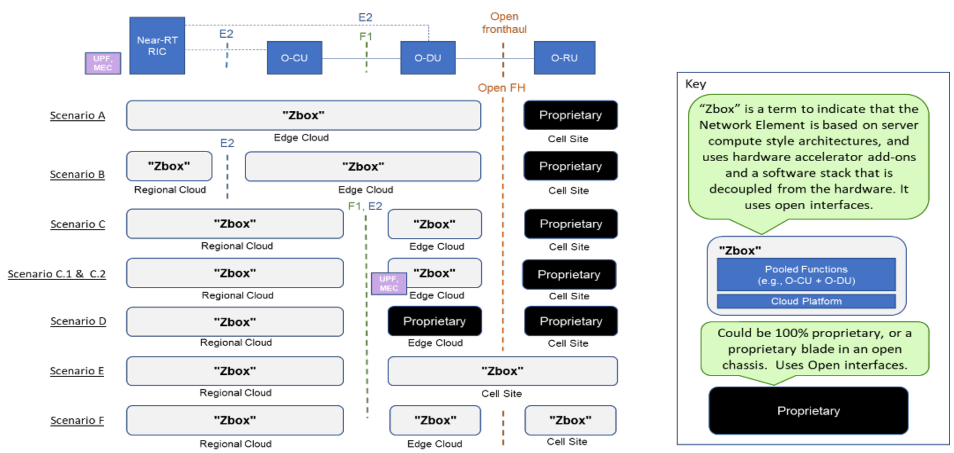
ML Model Lifecycle

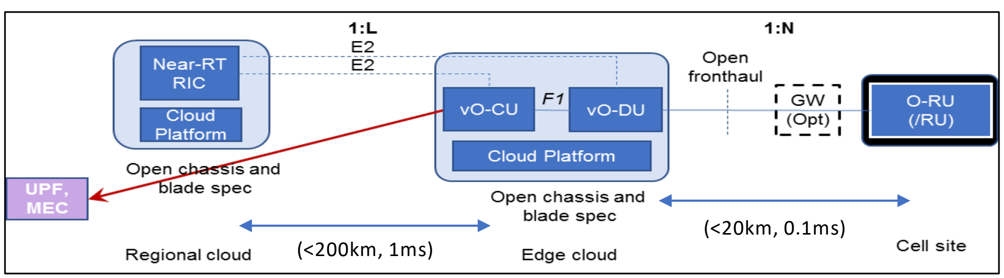


* ML Training Host can be part of non-RT RIC or can be executed offline

(Source ORAN)

Deployment models





Deployment Models

- Different models of deployment for Neal-Realtime RIC and CU/DU/RU
- Choice of deployment depends no user context, workload, and QoS requirements

Enterprise Scenario

- Specific deployment scenario converging outdoor and Pico cell deployment
- Near-RT RIC managing many CU-DU pairs
- Multiple RU aggregating to and Edge Cloud (Central Office)
- Ideal for Latency sensitive services

Deployment Example

- Outdoor mix of small and macro cells (mmWAVE + Sub 6Ghz)
- Limited fronthaul Capacity with MEC Requirement
- RIC in the Regional cloud and CU/DU in the Edge Cloud
- Cardinality
  - CU-DU :- 1:1
  - RIC – (CU-DU) :- 1:100+
  - DU-RU:- 1:32-64 TRPs

Fig. Source ORAN WG6

ORAN Use Cases

Use Cases within ORAN (1)

Use Case	Objective	Class	Impacted WGs	Impacted I/Fs	Impacted Entities
Context Based Dynamic Handover Management for V2X	Using AI/ML based application deployed at RIC to predict, present and mitigate handover anomalies based on real-time mentoring data on traffic, radio and road conditions on a UE level, and to ensure continuous and reliable connectivity between V2X UE and V2X AS.	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG1,WG2 ,WG3	O1/A1/E2	Non-RT RIC, Near-RT RIC, O-CU
Flight Path Based Dynamic UAV Resource Allocation	To perform radio resource allocation for on-demand coverage for UAV by ML model in RIC using the UE performance and application information such as flight path information, climate information, flight forbidden/limitation area information and Space Load information, etc.	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG1,WG2 ,WG3	O1/A1/E2	Non-RT RIC , Near-RT RIC, O-CU
Radio Resource Allocation for UAV Applications	In the UAV control vehicle scenario, to achieve the UE-level radio resource configuration optimization through the delivery of policies from Non-RT RIC and transfer of configuration parameters to RAN by Near-RT RIC in response to resource requirement from the application server.	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG1,WG2 ,WG3	O1/A1/E2	Non-RT RIC , Near-RT RIC, O-CU
QoE Optimization	L3 optimization to meet the challenge of both bandwidth intensive and delay sensitive applications by application QoE prediction and QoE based close-loop network optimization in real-time to proactively allocate resource before QoE is degraded, with the goal to optimize the QoE while radio resources being utilized in an efficient way	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG1,WG2 ,WG3	O1/A1/E2	Non-RT RIC, Near-RT RIC, O-CU
Traffic Steering	L3 optimization and automation to effectively achieve optimal traffic distribution among cells to improve network efficiency and enhance user experience with UE-centric and intelligent approach and ORAN architecture, and thus to provide network agility and automation for RAN and ultimately save OPEX.	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG1,WG2 ,WG3	O1/A1/E2	Non-RT RIC, Near-RT RIC, O-CU
Massive MIMO Optimization	L1 optimization and automation to proactively and continuously improve cell-centric network QoS and/or user (group)-centric QoE in a multi-cell and, possibly, multi-vendor massive MIMO deployment area with multiple transmission/reception points, depending on specific operator-defined objectives. The optimization is performed at different times scales by adapting beam configuration and policies with ML models.	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG1,WG2 ,WG3	O1/A1/E2	Non-RT RIC, Near-RT RIC, O-CU, O-DU

Use Cases within ORAN (2)

Use Case	Objective	Class	Impacted WGs	Impacted I/Fs	Impacted Entities
RAN Sharing	To enable an operator to share virtual RAN (logic implementation of the O-DU and O-CU functionalities) with another operator, and remotely monitor performance, configure and control resources in an infrastructure (RAN) that is owned by another operator using ORAN open interfaces, to ultimately reduce the network deployment cost	Virtual RAN Network	WG1,WG2 ,WG3, WG6	O1/O2/E2	Non-RT RIC, Near-RT RIC, O-CU, O-DU
QoS Based Resource Optimization	L3 QoS based resource optimization policy to overcome the shortcoming of static planning and configuration by ensuring that at least certain prioritized users can reach a satisfactory level of QoS through prioritizing certain user or user group utilizing the same service in certain particular (congestion) situation with A1 QoS policy.	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG1,WG2 ,WG3	A1/O1/E2	Non-RT RIC, Near-RT RIC, O-CU
RAN Slice SLA Assurance	To ensure RAN slice SLA by executing ML model(s) in Non-RT RIC (slow loop) and/or Near-RT RIC (fast loop) using long term and/or ear-real-time RAN slice subnet PKIs, fine-tuning RAN behavior dynamically through O1 configuration, A1 policy and E2 control.	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG2,WG3 ,WG6	O1/O2/A1 /E2	Non-RT RIC, Near-RT RIC, O-CU, O-DU
Low Cost Radio Access Network White-box Hardware	L2 and L1 complete reference design for high performance, spectral and energy efficient white-box base stations to reduce CAPEX of 5G development & deployment. Focus on O-DU, O-RU and FH WG, Split option 7-2x supported, option 6 and 8 being considered.	White-box Hardware Design	WG7	Open fronthaul interface	O-DU, O-RU
Dynamic Spectrum Sharing between 4G and 5G	To provide an open, vendor-agnostic, centralized dynamic spectrum sharing mechanism for 4G and 5G systems to share limited critical bands in the existing 4G deployment based on traffic load dynamically using the open architecture of RIC, aiming the efficient and seamless evolution from 4G to 5G. ML models or algorithms for dynamic radio resource allocation between 4G and 5G are used in Non-RT RIC and/or Near-RT RIC with long and/or near-real-time control loop	AI-enabled RAN and Open RAN Interfaces (O1/A1/E2).	WG1,WG2 ,WG3	O1/A1/E2	Non-RT RIC , Near-RT RIC, O-CU. O-DU

Use Cases Motivating ORAN

- **L3 : Service Optimization (User Performance or Spectral Efficiency)**
  - Traffic Steering
    - Move from cell-centric to Use Centric Optimization
  - QoE Optimization
    - Application classification, QoE prediction, and available bandwidth prediction
- **L2 :RAN Service Assurance (Application Specific Optimization)**
  - Dynamic RAN Slicing and Management
  - Dynamics Spectrum Sharing
- **L1 : RAN Automation (Minimizing Human Effort)**
  - 3D MIMO Optimization
    - Adaptive Beam Configuration in non-Realtime and Real time conditions

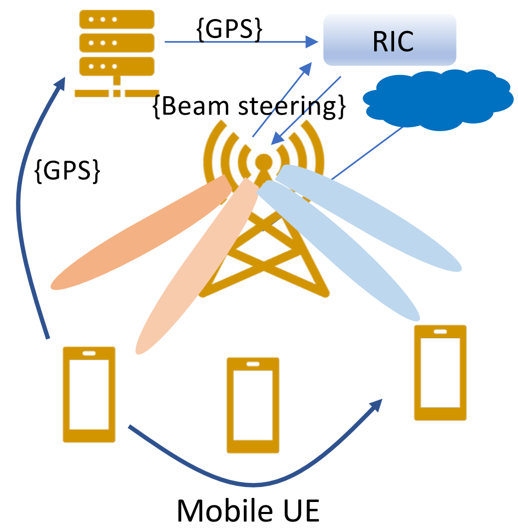
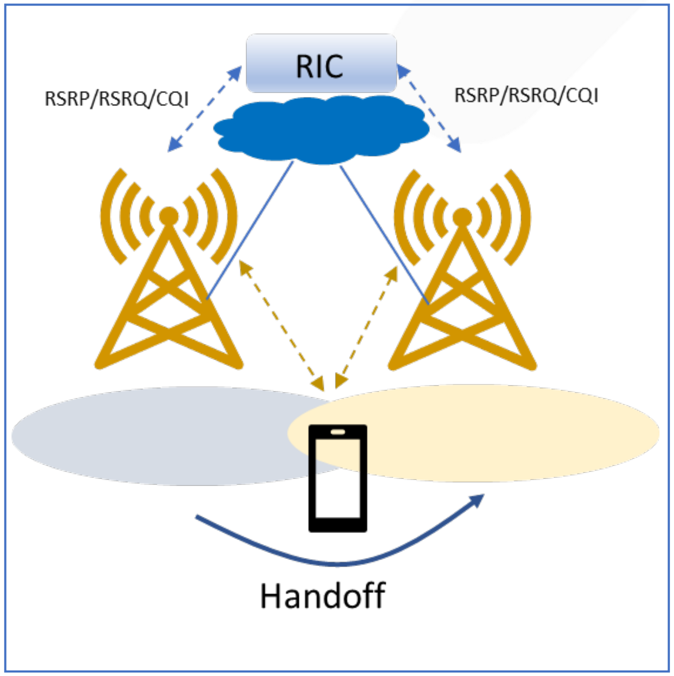


Fig. Source ORAN WG1

1. Traffic Steering
2. Dynamic Spectrum Sharing

Traffic Steering and Mobile Load Balancing

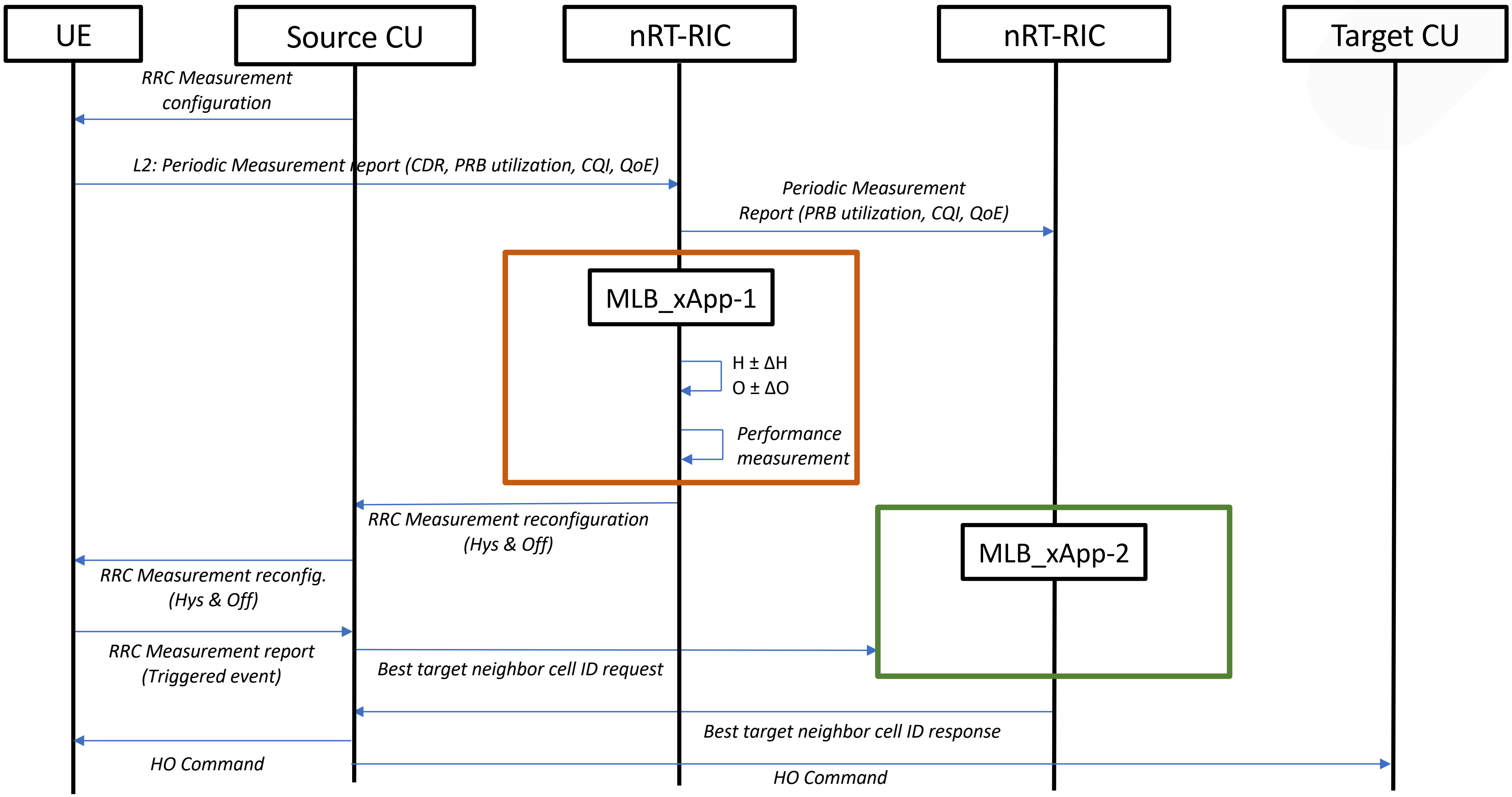
- **Objectives**
  - Balancing load across cells within a region, e.g. within Central Office
  - To ensure the load is transferred to cells with sufficient residual capacity
  - MLB can be also be triggered due to user centric performance metric
- **Non-Realtime and Realtime RIC View**
  - Intents
    - Trigger based intents such as cell load above certain threshold, offload traffic threshold doesn't exceed beyond another threshold
  - Real time Control
    - Performance metrics subscription and report
      - Call_drop_ratio, DL_prb_utilization, UE_CQI, UE_RSRP
    - Affecting hand-over to cell (with priority of certain cells if required) with better RSRP, or within the same cell allocating more PRBs



- **Algorithms**
  - AI/ML models trained on O1 data based on cell-centric and UE-centric parameters [1]
  - Traditional RRM techniques based on Hysteresis and Offset manipulation to trigger A1-A6 condition in UEs when cell load or related parameter threshold are violated. [2]

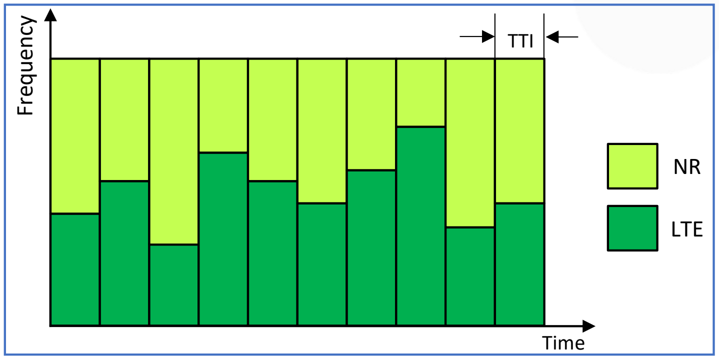
[1] ATT ORAN WG2 contribution on AI/ML use case for Traffic Steering.
[2] Wen Yu Li et al, "A Novel Dynamic Adjusting Algorithm for Load Balancing and Handover Co-Optimization in LTE", Journal of Computer Science and Technology, 2013

Workflow for MLB



Dynamic Spectrum Sharing

- While new C-band and mmWave spectrum offers higher bandwidth for 5G, it suffers from path and penetration loss limiting its coverage
- 5G coverage at a national scale would require
  - Dense BS deployment using the high bands
    - Cost and ICI challenges
  - Re-farming LTE spectrum to 5g
    - May affect LTE performance and underutilization of re-farmed resources
  - Use DSS techniques to dynamically share the low band spectrum
- Benefit of DSS: cost effective and efficient way to achieve coverage and help easy rollout and seamless migration to 5G

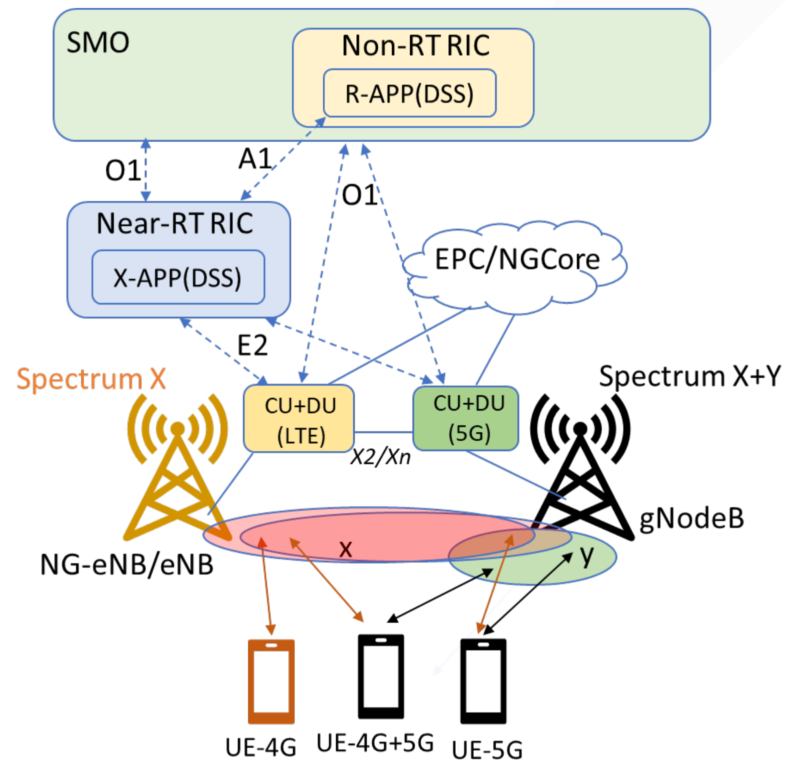


This use case focusses on enabling DSS configuration, management and near-real-time control of the CU/DU nodes participating in DSS over the ORAN architecture.

Architecture of DSS over RIC Framework

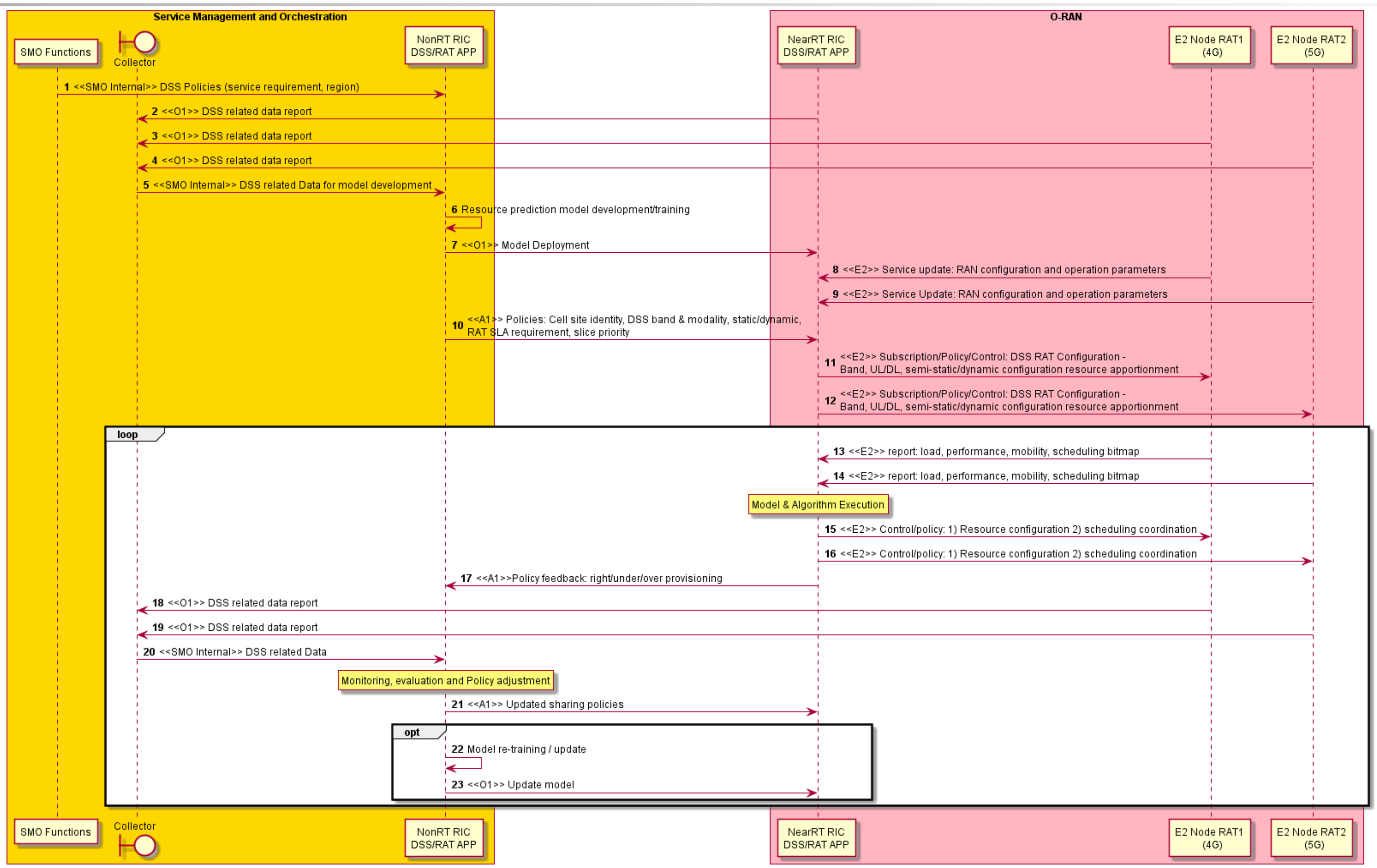
Benefits

- **Open architecture and interfaces**
  - Resource allocation and interference mitigation algorithms, interoperability for BSs from different vendors
- **RAN data analytics and AI/ML framework**
  - Prediction of traffic, mobility events, flash crowd scenarios for efficient resource sharing
- **Using global knowledge**
  - Fine granular information (e.g. mobility) of 4G/5G users to pre-provision neighboring 4G/5G cells with appropriate resource split which is not possible in a local context
- **Utilizing Near Realtime Functions**
  - Interference Mitigation: Synchronize 4G/5G schedulers to enable orthogonal scheduling
    - Either use a logical centralized function over proprietary interfaces, or leverage E2 for such synchronization considering NearRT RIC is collocated and engineered for ~ms latency
    - Could in some cases also avoid the use of X2/Xn interface which otherwise may incur additional latency



Architecture of DSS over RIC Framework

End-to-end DSS Workflow



Conclusions

- OpenRAN is happening and is being deployed
- Overall Cost of these networks will be much lower, with high softwarization of the management, control and user plane infrastructure.
- Data driven RRM optimization using new functions such as RIC over open interfaces is being driven by forums like ORAN
- Operators get to choose best of the breed RRM Applications or build on their own

Thank You!