openrangym.com

Institute for the Wireless Internet of Things at Northeastern University

Intelligent networks with Open RAN
Challenges and opportunities

Michele Polese

Institute for the Wireless Internet of Things

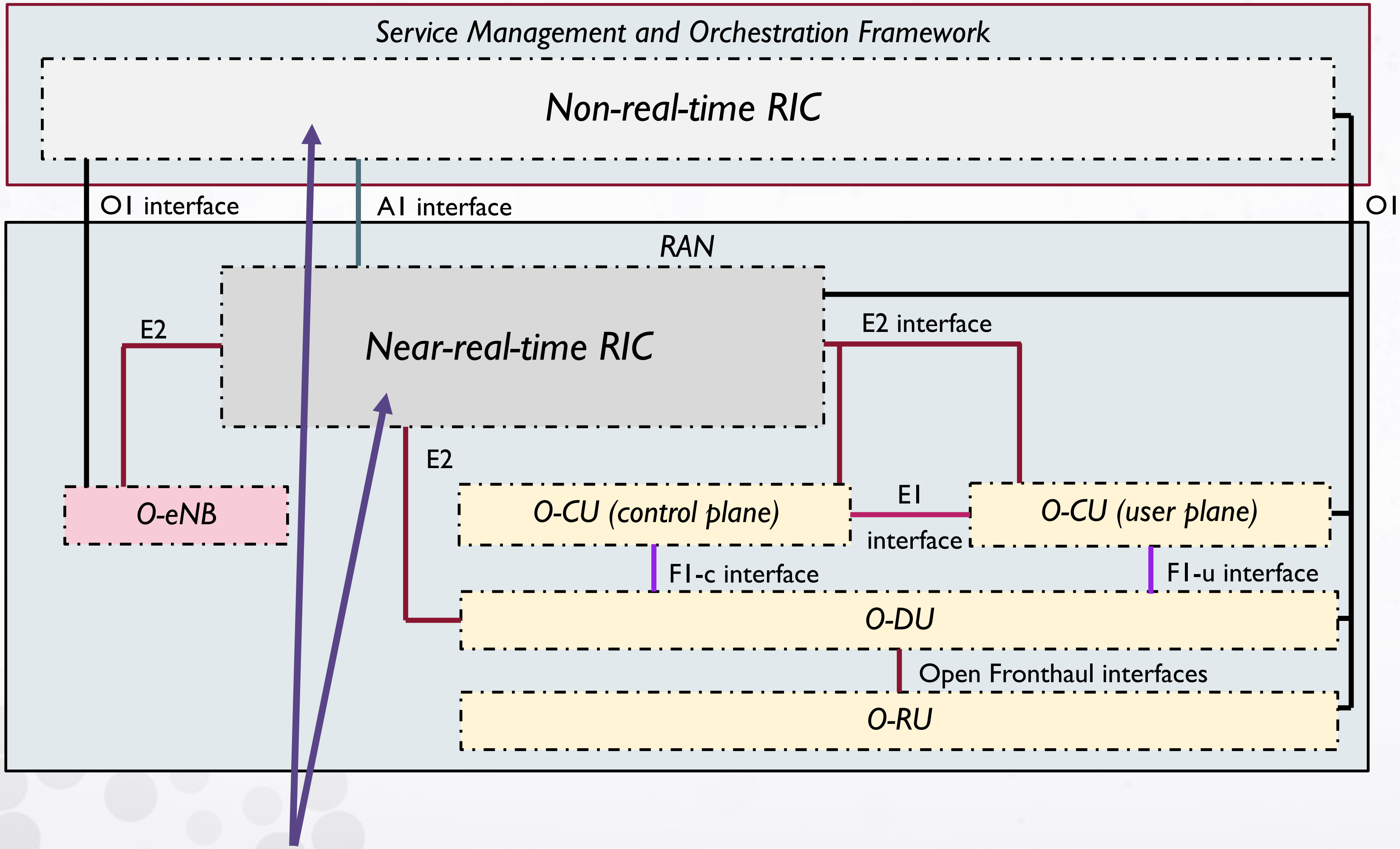
Northeastern University

m.polese@northeastern.edu

with Leonardo Bonati, Salvatore D'Oro, Stefano Basagni, Tommaso Melodia

Partially supported by NSF Grants CNS-1925601, CNS-2120447, and CNS-2112471 and ONR Grant N00014-20-1-2132

O-RAN ALLIANCE - a reference architecture for programmable NextG



1. Open, standardized interfaces
2. Disaggregated RAN
3. Open-source contributions

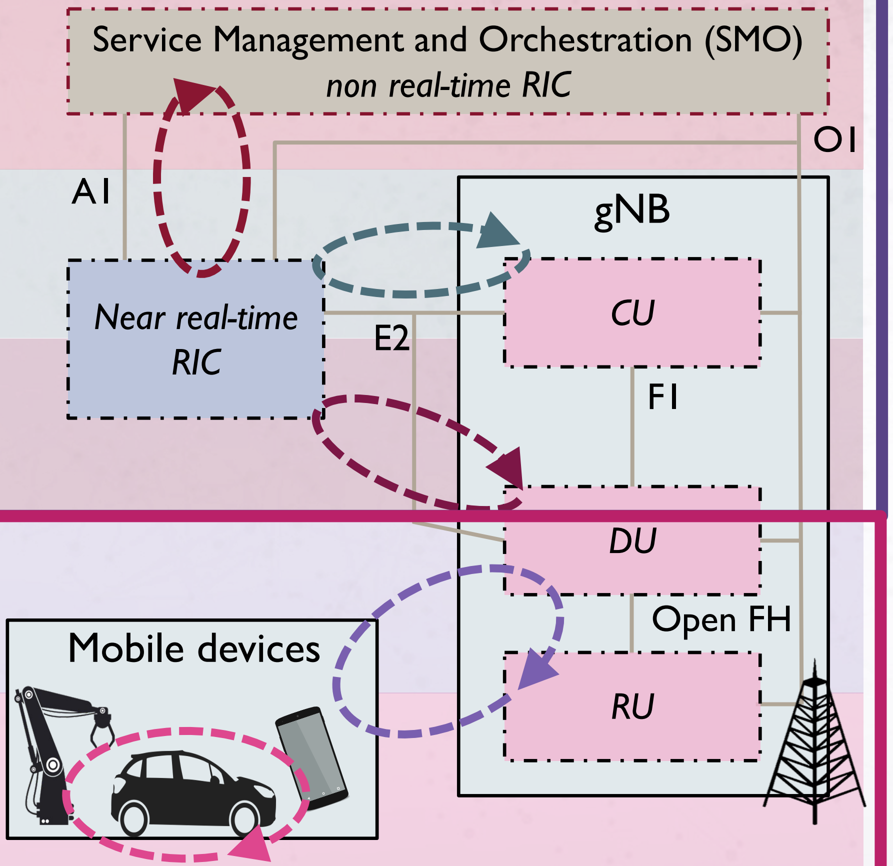
4. RAN Intelligent Controllers

Institute for the Wireless Internet of Things at Northeastern

Intelligent Control Loops

Currently supported by O-RAN

Control and learning objective	Scale	Input data	Timescale	Architecture
Policies, models, slicing	> 1000 devices	Infrastructure-level KPIs	Non real-time > 1 s	
User Session Management e.g., load balancing, handover	> 100 devices	CU-level KPIs e.g., number of sessions, PDCP traffic	Near real-time 10-1000 ms	
Medium Access Management e.g., scheduling policy, RAN slicing	> 100 devices	MAC-level KPIs e.g., PRB utilization, buffering	Near real-time 10-1000 ms	
Radio Management e.g., resource scheduling, beamforming	~10 devices	MAC/PHY-level KPIs e.g., PRB utilization, channel estimation	Real-time < 10 ms	
Device DL/UL Management e.g., modulation, interference, blockage detection	1 device	I/Q samples	Real-time < 1 ms	



For further study or not supported

Institute for the Wireless Internet of Things at Northeastern

Open Challenges toward Intelligent Open RAN

- Need large-scale heterogeneous datasets
- Need testing of closed-loop control without compromising network performance
- Need algorithms that generalize to different scenarios and conditions

Experimental platforms for wireless AI

Need large-scale heterogeneous datasets

PAWR platforms and Colosseum can be used
*to collect **datasets** at scale*

Tools are available for large-scale data collection in cellular networks: SCOPE platform https://github.com/wineslab/colosseum-scope

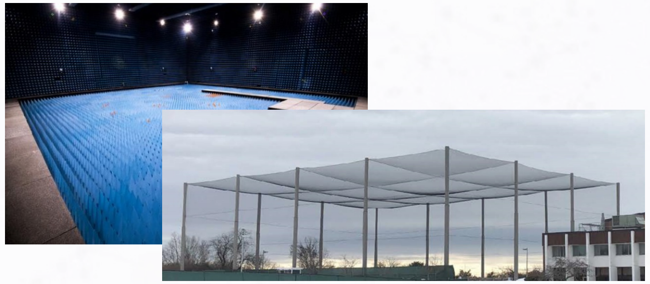
https://advancedwireless.org
https://northeastern.edu/colosseum/

Experimental Research at Northeastern

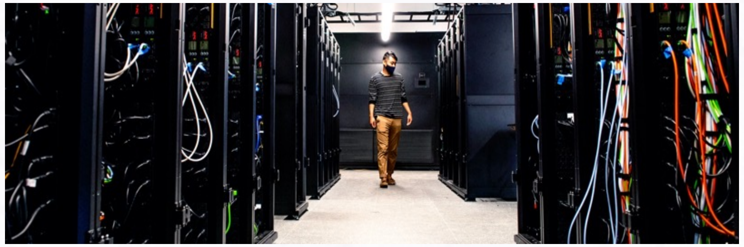
*Develop and validate **innovative spectrum solutions** in heterogeneous environments*

Arena @ NEU

UAS Lab @ NEU



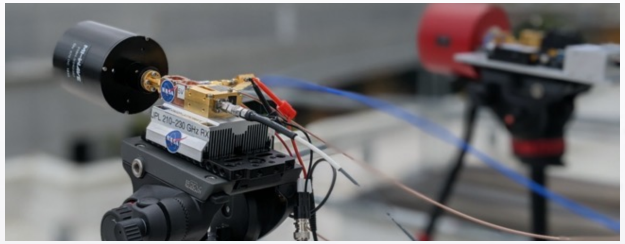
Colosseum @ NEU



X-Mili @ NEU

5G O-RAN ALLIANCE

mmWave/THz @ NEU



PAWR Platforms

POWDER

Salt Lake City, UT
Software defined networks and massive MIMO

COSMOS

West Harlem, NY
Millimeter wave and backhaul research

AERPAW

Raleigh, NC

Rural Broadband Platform

TBD
Coming late 2020

An Experiment's Journey

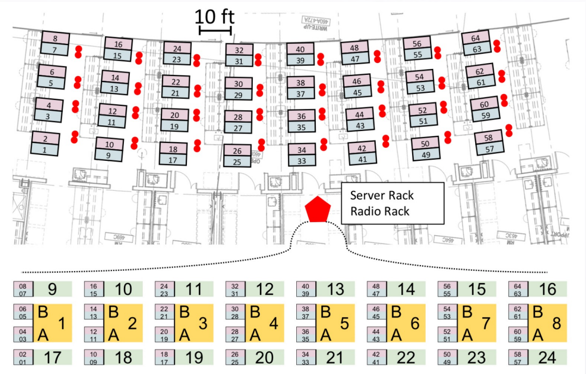
The same experiment (and software) can be seamlessly deployed in the different testbeds

- Initial design and testing at-a-scale on Colosseum w/ different scenarios
- Validate on real-world indoor environment on Arena
- Experiment into the wild on city-scale platforms

Test at-a-scale on emulated scenarios

COLOSSEUM at Northeastern University

Validate in real wireless environment



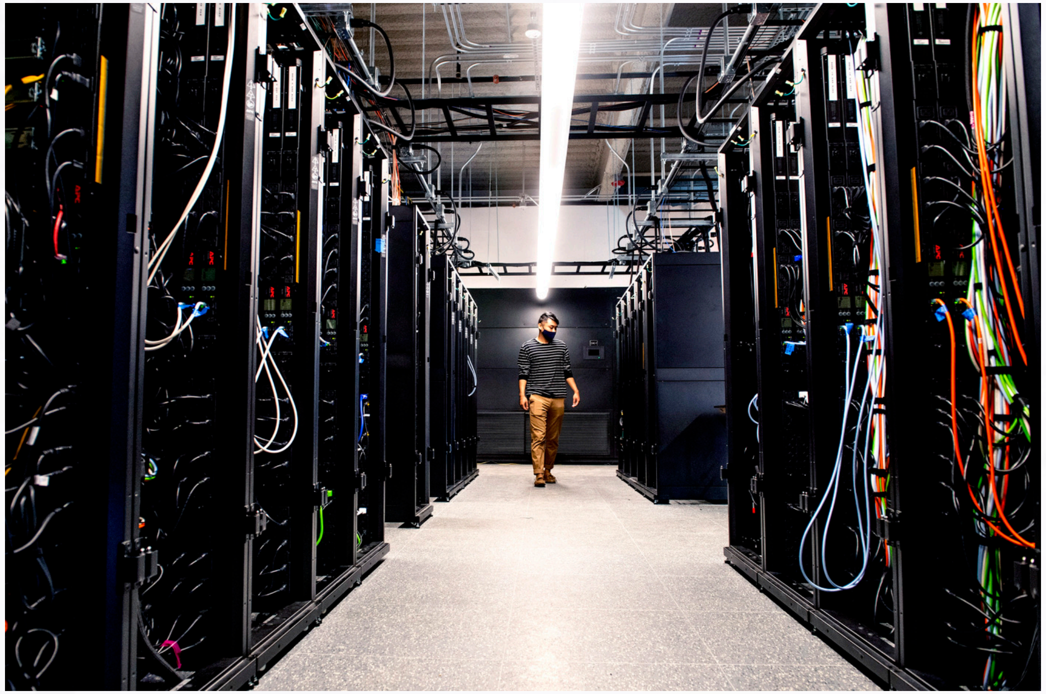
Test large-scale capabilities

Platforms for Advanced Wireless Research

AERPAW

Colosseum @ Northeastern

Colosseum is the world most powerful hardware-in-the-loop network emulator

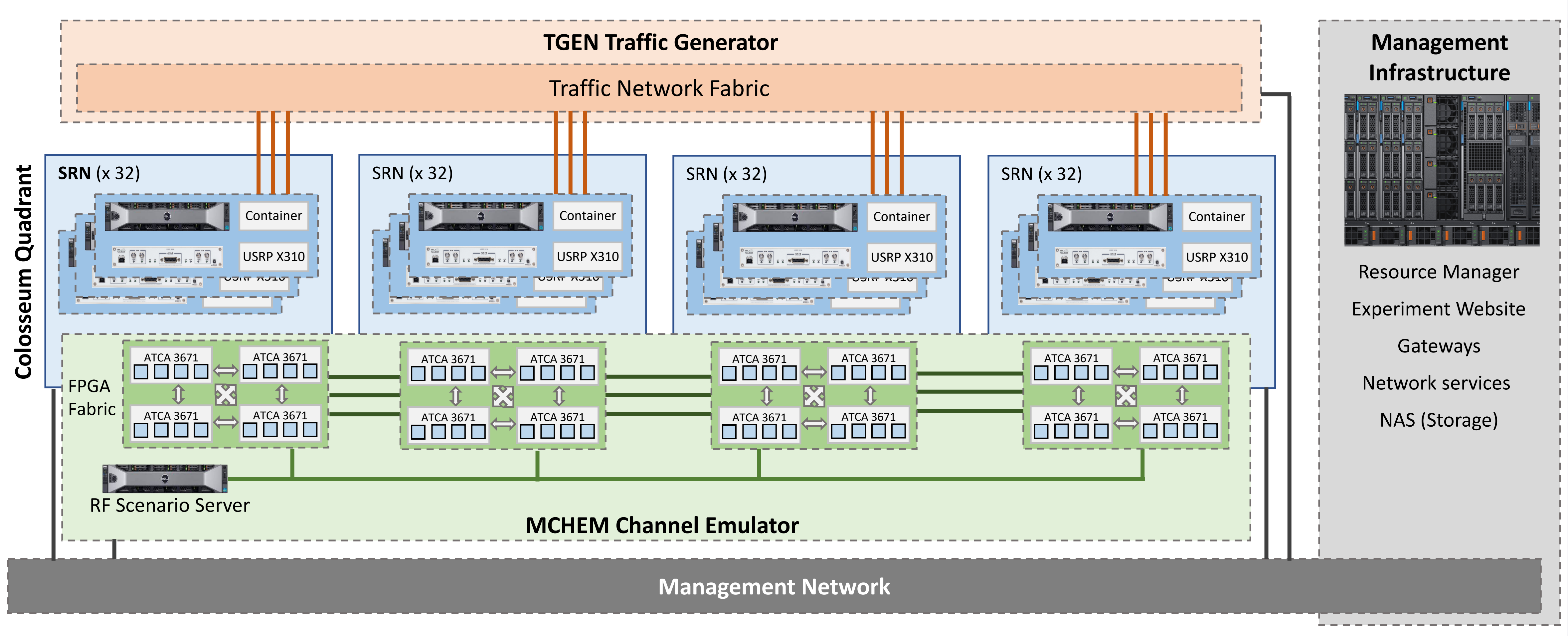


- 256 software-defined radios
- 25.6 GHz of emulated bandwidth, 52 TB/s RF data
- 21 racks of radios, 171 high-performance servers w/ CPUs, GPUs
- Massive computing capabilities (CPU, GPU, FPGA):
  - > 900 TB of storage
  - 320 FPGAs
  - 18 10G switches
  - 19 clock distribution systems
  - 52 TB/s of digital RF data

Create and test complex 5G scenarios

Institute for the Wireless Internet of Things at Northeastern

Colosseum Architecture



Open Challenges

Need testing of closed-loop control without compromising network performance

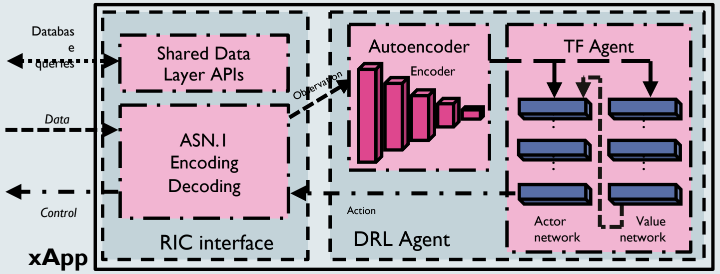
OpenRAN Gym

A toolbox for Intelligent O-RAN

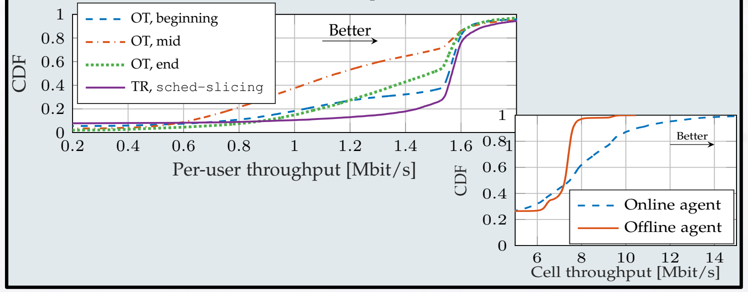
www.openrangym.com

Collect datasets at scale on virtual RF scenarios

Design, train, and package ML solutions for O-RAN



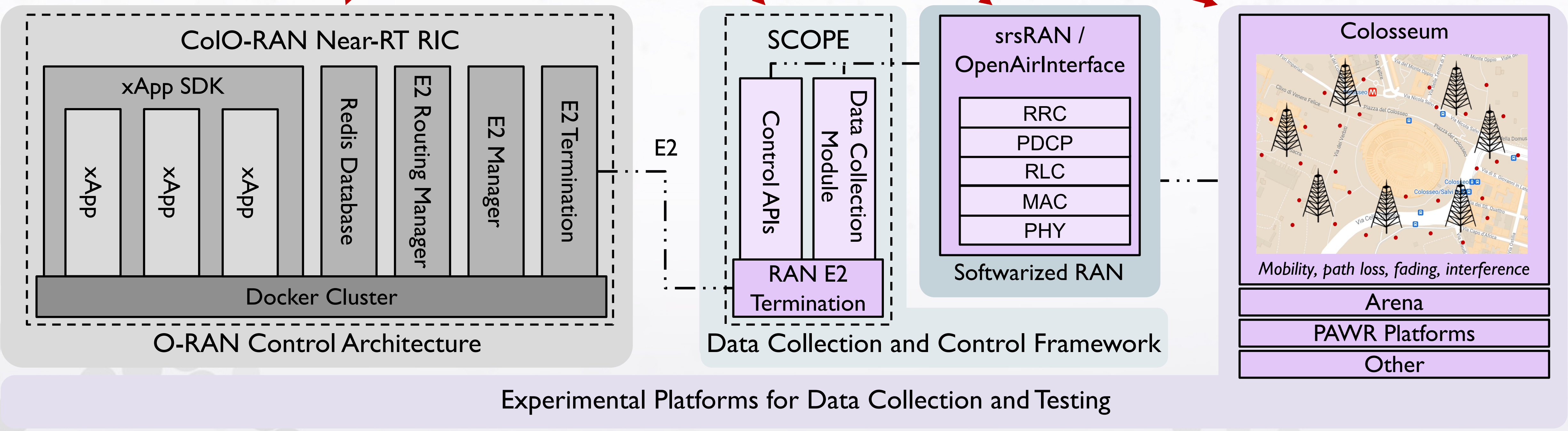
Test and refine on experimental wireless platforms



L. Bonati, M. Polese, S. D'Oro, S. Basagni, and T. Melodia, *"OpenRAN Gym: An Open Toolbox for Data Collection and Experimentation with AI in O-RAN,"* in Proc. IEEE WCNC Workshops, 2022

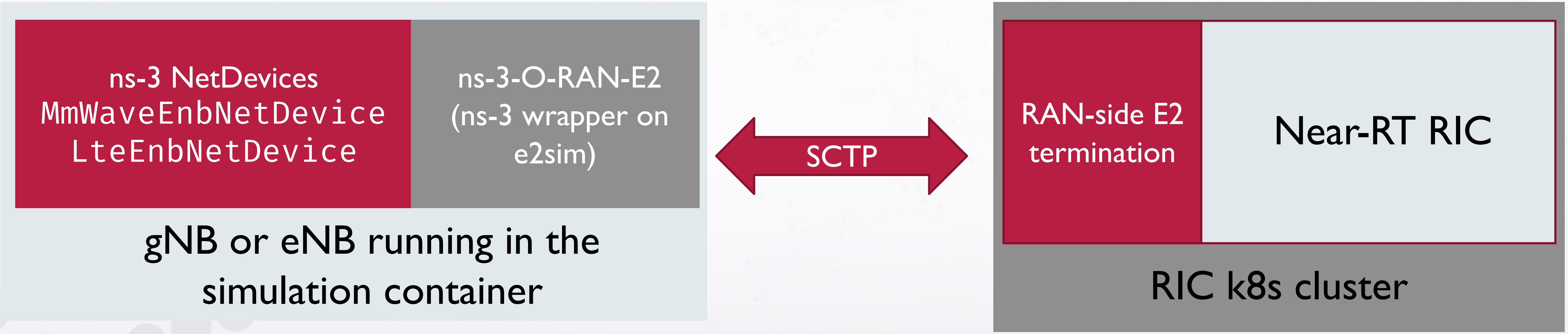
OpenRAN Gym – A Toolbox for Intelligent O-RAN

- O-RAN-compliant **near-real-time RIC** running on Colosseum (ColO-RAN)
- RAN framework for **data-collection and control** of the base stations (SCOPE)
- **Programmable** protocol stacks (based on srsRAN at this time)
- Publicly-accessible **experimental platforms** (e.g., Colosseum, Arena, PAWR platforms)



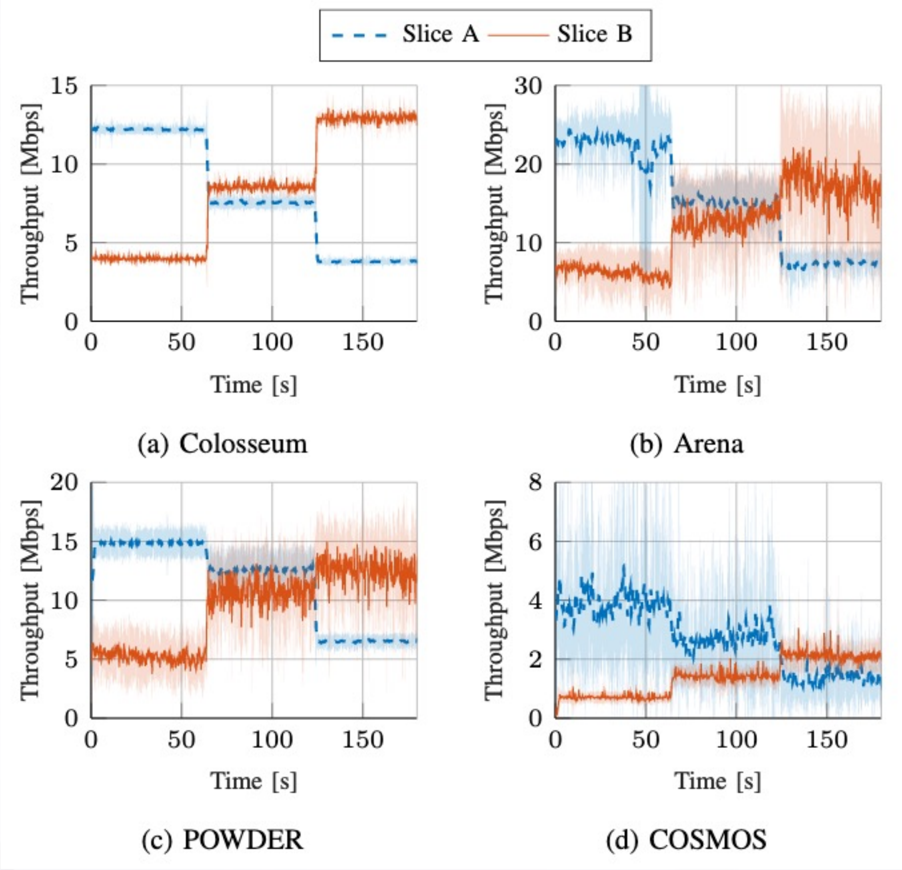
OpenRAN Gym and ns-3

- Developed a custom E2 termination for ns-3
- ns-3 provides functional RAN environment and connects to an O-RAN-compliant near-RT RIC
- To be included in the O-RAN SC

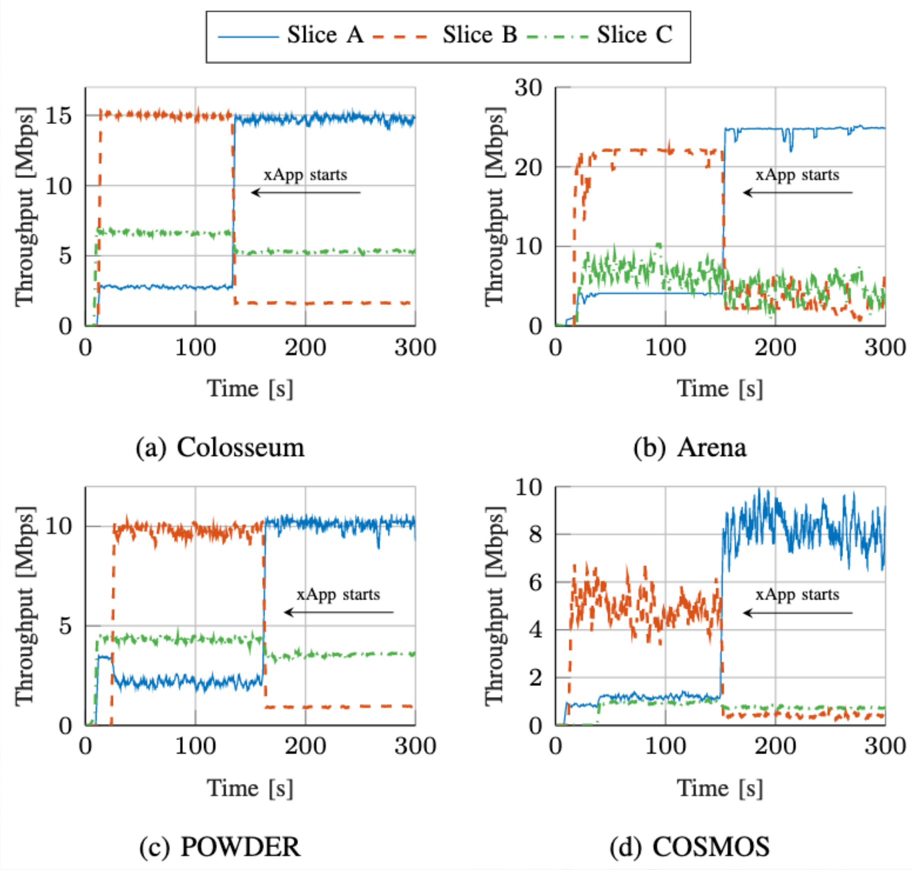


OpenRAN Gym on PAWR Platforms

Periodic change of slicing resources



xApp closed-control loop



*Results are consistent across **very different platforms** with **heterogeneous environments***

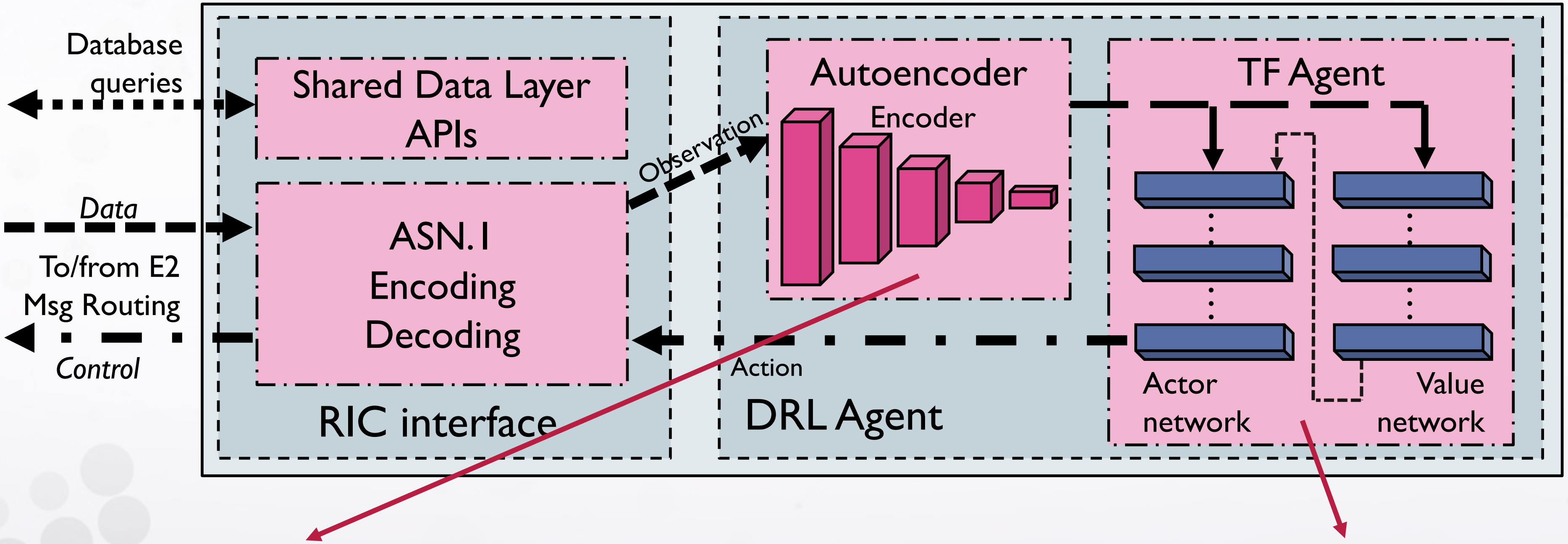
L. Bonati, M. Polese, S. D'Oro, S. Basagni, T. Melodia, "*OpenRAN Gym: AI/ML Development, Data Collection, and Testing for O-RAN on PAWR Platforms,*" arXiv:2207.12362 [cs.NI], July 2022.

Institute for the Wireless Internet of Things at Northeastern

ColO-RAN – ML development and testing for O-RAN

Need algorithms that generalize to different scenarios and conditions

ColO-RAN xApp



Generate a compressed representation of the RAN

Exploit it to generate control actions in the network

M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "ColO-RAN: Developing Machine Learning-based xApps for Open RAN Closed-loop Control on Programmable Experimental Platforms", IEEE TMC, to appear

ColO-RAN xApps

Control slicing and scheduling of the RAN

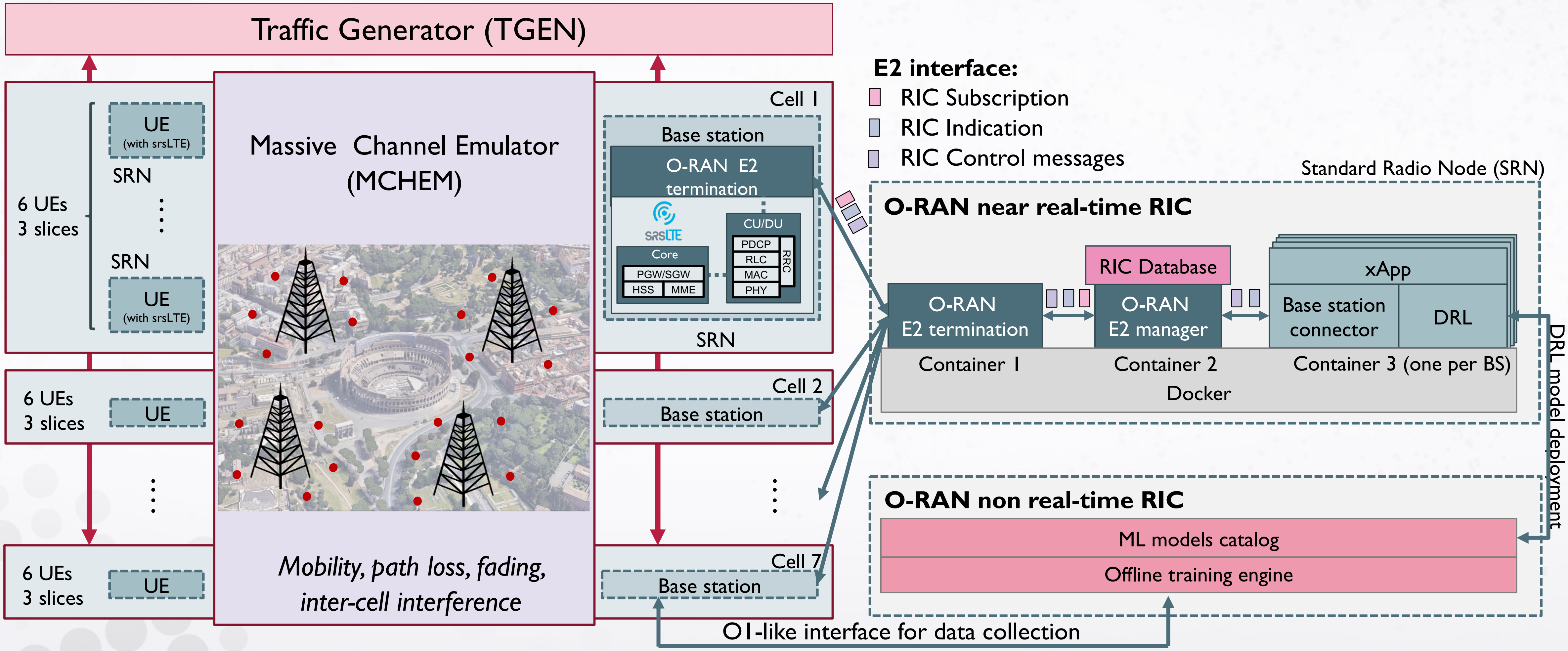
- `sched` xApp – select scheduling policy for a specific slice
- `sched-slicing` – jointly select scheduling policy and slicing for a base station

xApp	Functionality	Input (Observation)	Output (Action)	ML Models	Utility (Reward)
`sched-slicing`	Single-DRL-agent for joint slicing and scheduling control	Rate, buffer size, PHY TBs (DL)	PRB and scheduling policy for each slice	DRL-base, DRL-reduced-actions, DRL-no-autoencoder	Maximize rate for eMBB, PHY TBs for MTC, minimize buffer size for URLLC
`sched`	Multi-DRL-agent per-slice scheduling policy selection	Rate, buffer size, PRB ratio (DL)	Scheduling policy for each slice	DRL-sched	Maximize rate for eMBB and MTC, PRB ratio for URLLC

Institute for the Wireless Internet of Things at Northeastern

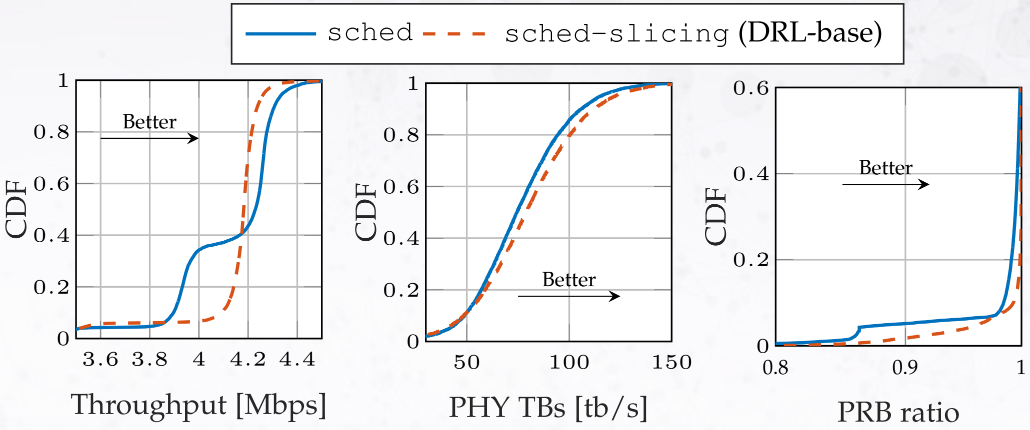
ColO-RAN Testing Deployment – 42 users and 7 base stations

OpenRAN Gym on a large-scale Colosseum deployment – 7 base stations, 42 UEs, 3 slices

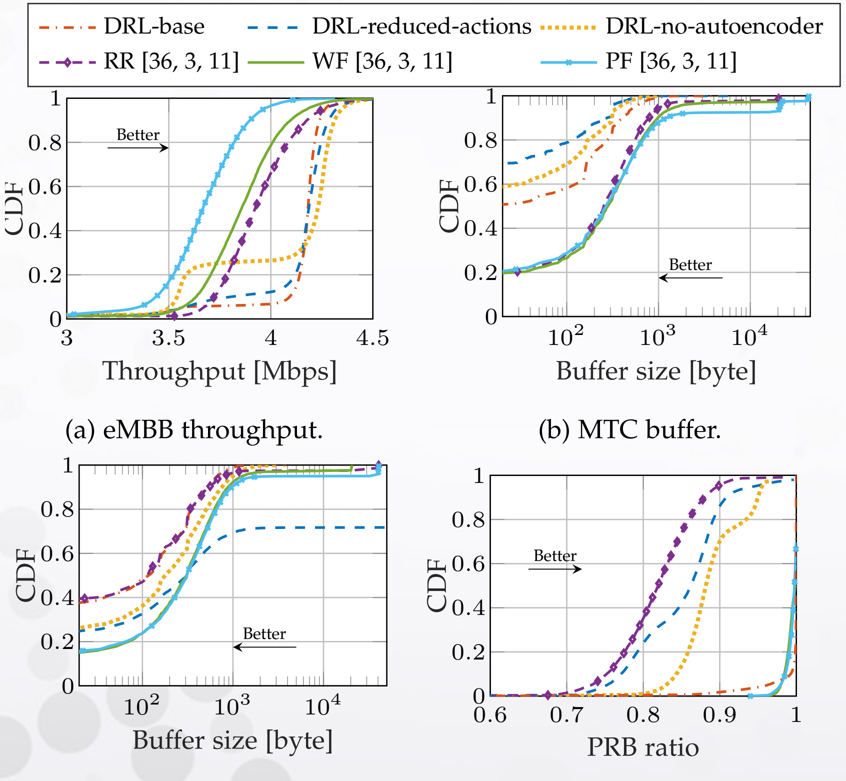


ColO-RAN results

Joint control (scheduling and slicing) outperforms scheduling-only control *for all slices*



(a) eMBB throughput. (b) MTC PHY TBs. (c) URLLC PRB ratio.



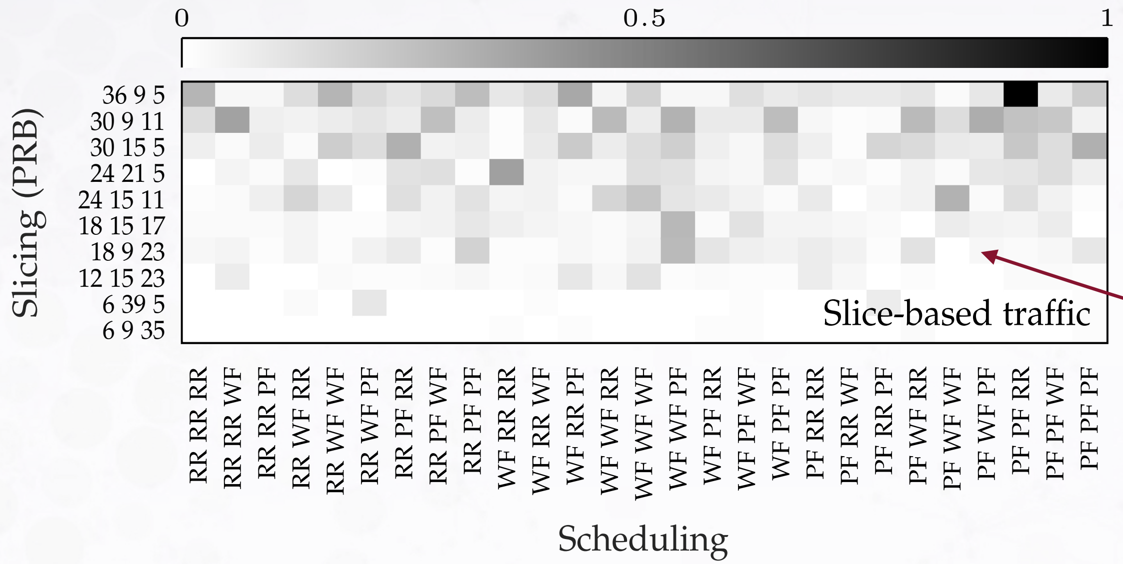
(a) eMBB throughput. (b) MTC buffer.

(c) URLLC buffer. (d) URLLC PRB ratio.

Best performance from proper action space design + autoencoder

Institute for the Wireless Internet of Things at Northeastern

ColO-RAN – online training



The agent is trained with a dataset that represent certain network conditions

As a result, it selects optimal network configurations with a certain probability

What happens when there is an unforeseen configuration in the network?

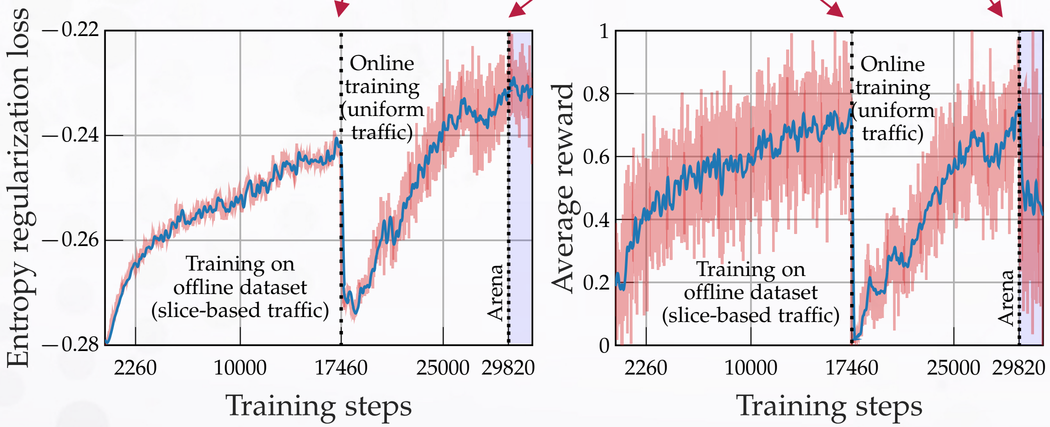
Fine-tune the DRL model with online training on the near-RT RIC itself

Institute for the Wireless Internet of Things at Northeastern

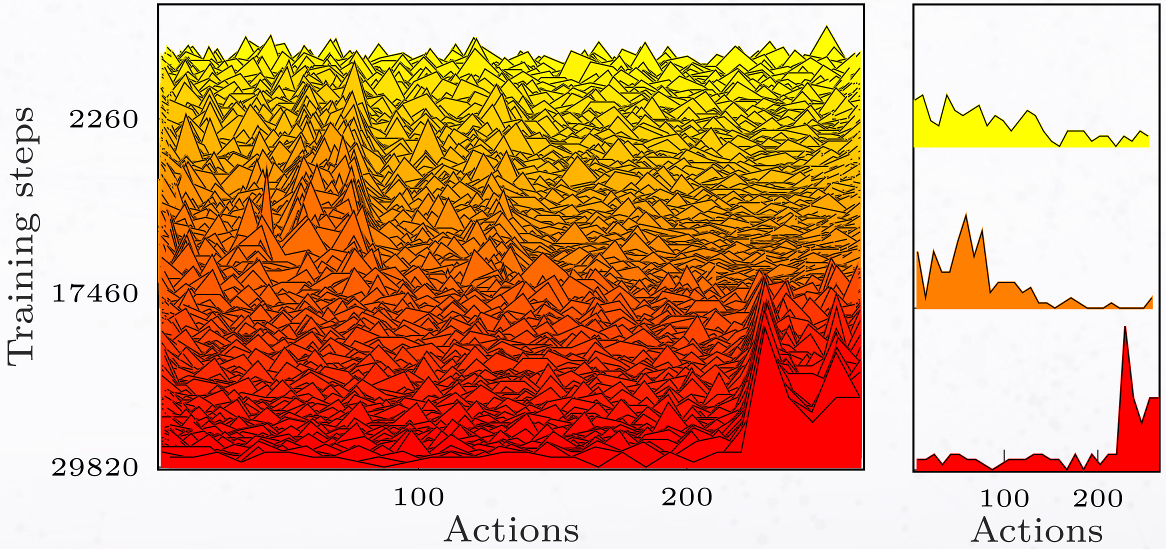
ColO-RAN – online training

Two events:

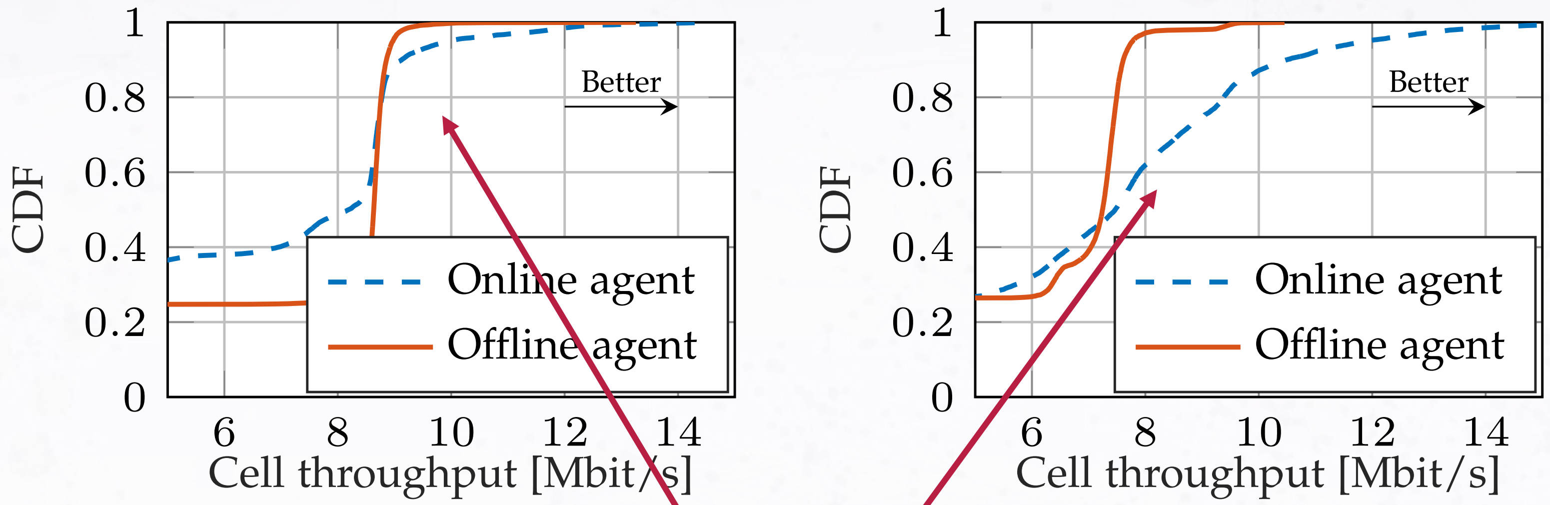
1. change user traffic (from slice-based to uniform)
2. move from Colosseum to Arena deployment



(a) Entropy regularization loss. (b) Reward.



ColO-RAN – online training



(a) Slice-based source traffic. (b) Uniform source traffic.

The online-trained agent

- adapts to the new traffic
- still remember how to behave with the old traffic

Institute for the Wireless Internet of Things at Northeastern

Orchestrating RAN Intelligence

Four open problems:

High-level intent

I want to change scheduling decisions in real time to broadcast 4K video in Times Square, NY

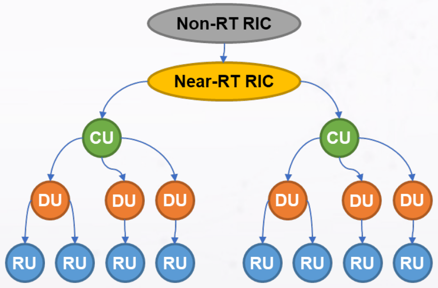
1. What does the intent mean?

ML/AI catalog

2. Which models best satisfy the intent?

3. Where and when to deploy network intelligence?

4. How to deploy, execute, manage intelligence?



Infrastructure

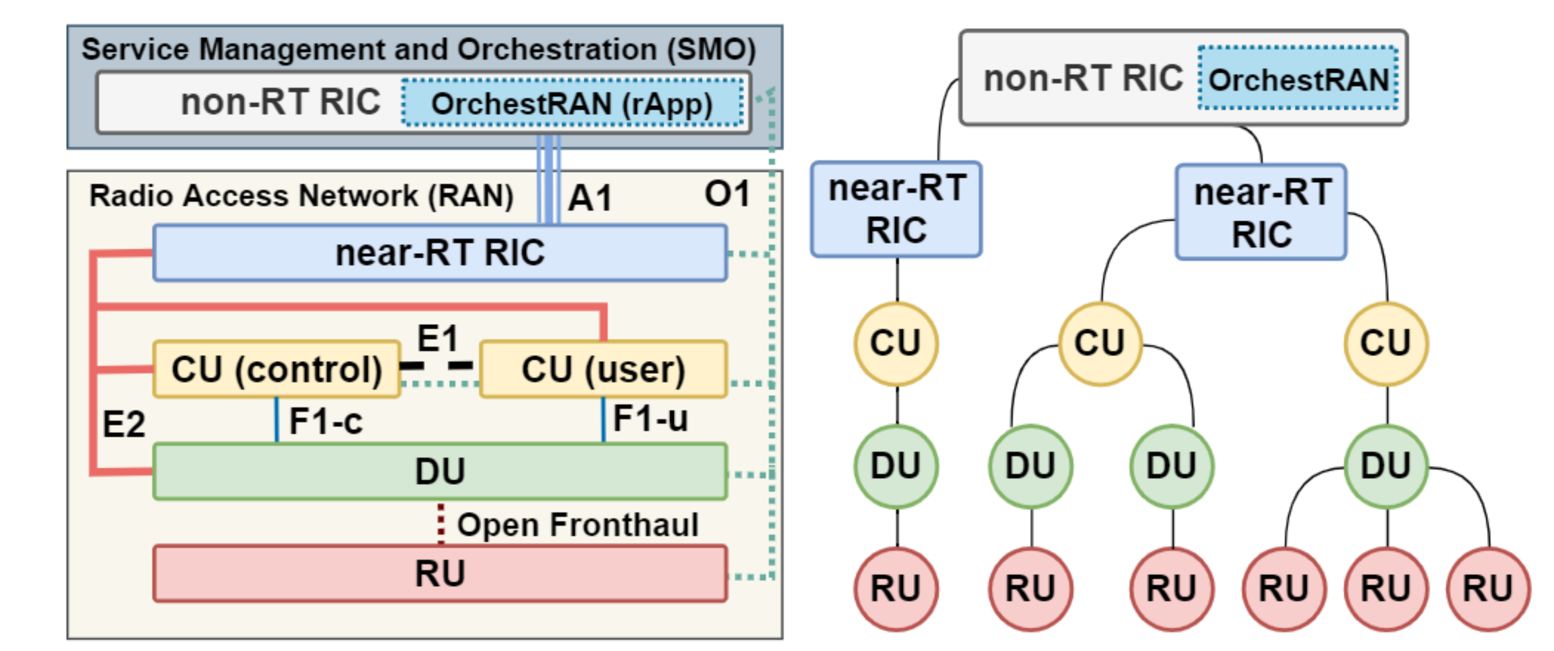
OrchestRAN: orchestrating intelligence in the Open RAN

Execute in the non-RT RIC

I. **Intent** recognition

II. **Optimized** intelligence **placement**

III. **Automated** deployment/execution/management of intelligence



S. D'Oro, L. Bonati, M. Polese, T. Melodia, "OrchestRAN: Network Automation through Orchestrated Intelligence in the Open RAN," Proc. of IEEE Intl. Conf. on Computer Communications (INFOCOM), May 2022.

OrchestRAN step-by-step

(II) Collect requests

(III) Compute orchestration policy

--- Which model

--- Where and when

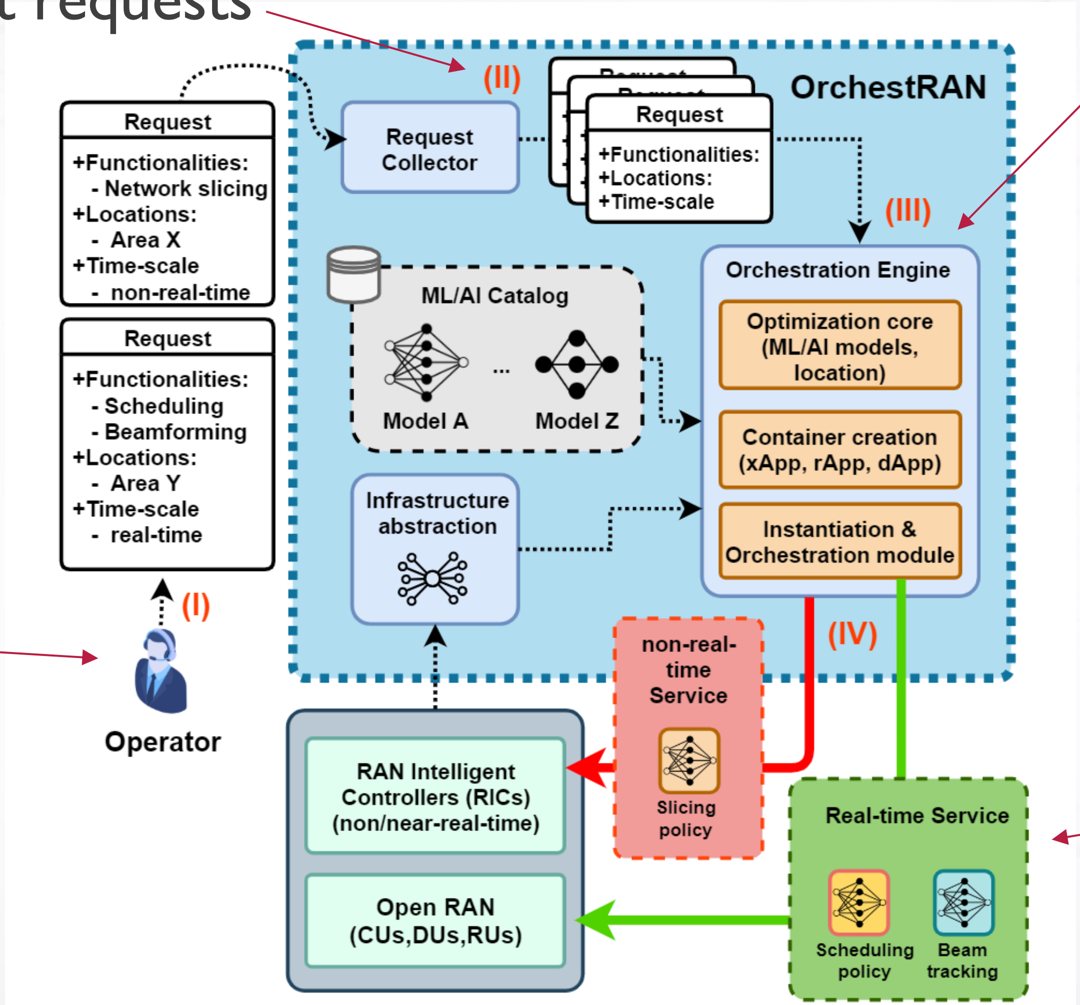
(I) Submit request:

--- Functionalities

--- Locations

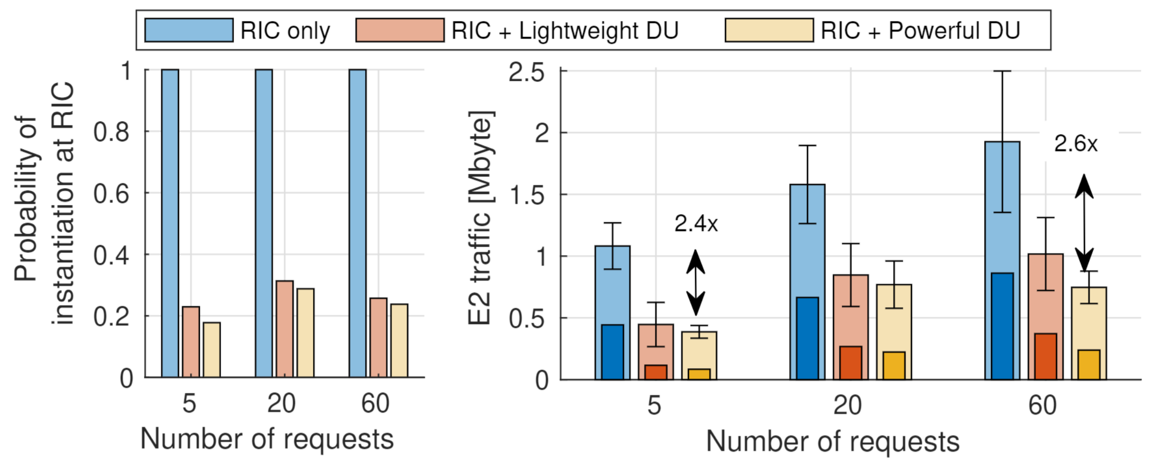
--- Time-scale

(IV) Deploy intelligence



Prototyping OrchestRAN

- Orchestrate the ColO-RAN xApps on Colosseum
  - **E2 traffic:**
    - **Light bars:** total traffic
    - **Dark bars:** payload only
  - **xApps** on near-RT RIC
  - **dApps** on DUs
    - These are not O-RAN-compliant (yet)



More powerful Dus
More intelligence at the edge

Only 40% is payload

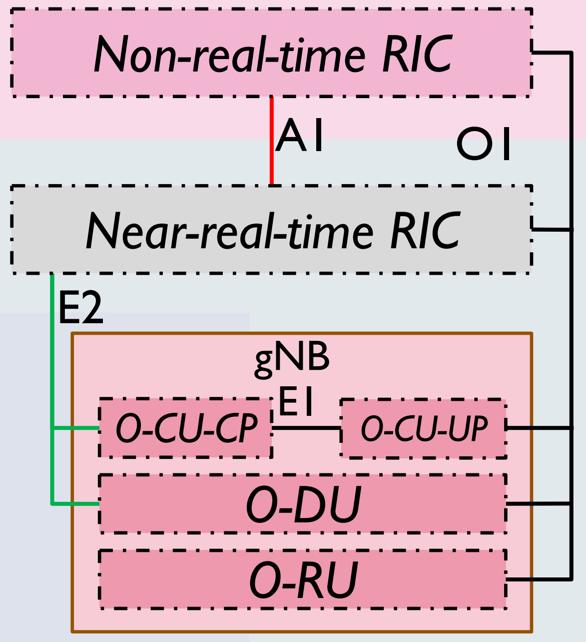
Institute for the Wireless Internet of Things at Northeastern

dApps – the missing piece in O-RAN

Control objective	Input data	Timescale and Apps
Policies, models, slicing	Aggregated KPMs	Non-real-time > 1 s **rApps**
Radio Resource Management, Session Management	CU-level and MAC-level KPMs	Near-real-time 10-1000 ms **xApps**
Beamforming, Scheduling, Puncturing, Interference and Modulation Management	MAC/PHY-level KPMs, I/Q samples, Packets	Real-time < 10 ms **dApps**

dApps

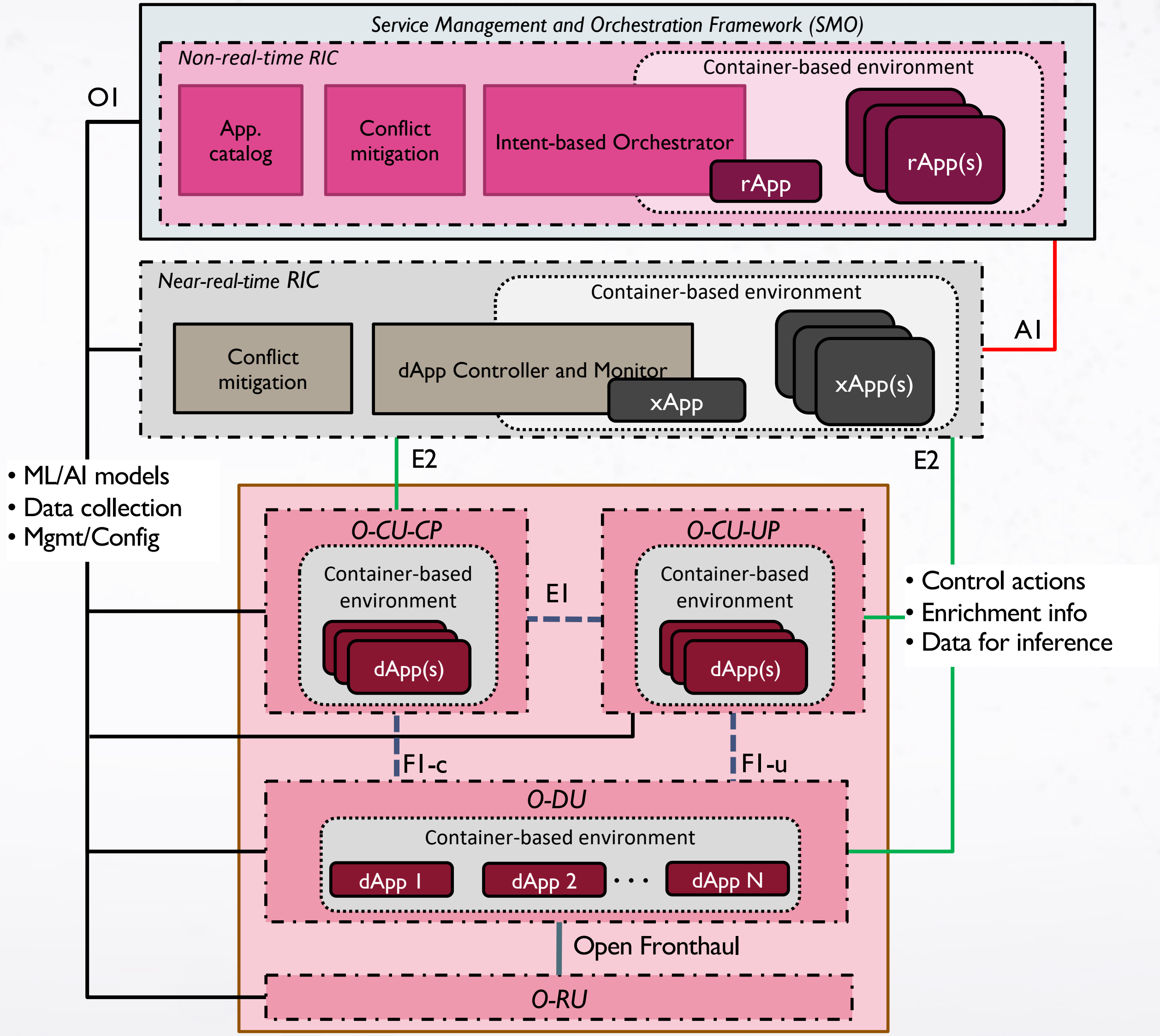
O-RAN Architecture



S. D'Oro, M. Polese, L. Bonati, H. Cheng, and T. Melodia, "dApps: Distributed Applications for Real-time Inference and Control in O-RAN," arXiv:2203.02370 [cs.NI], March 2022.

Institute for the Wireless Internet of Things at Northeastern

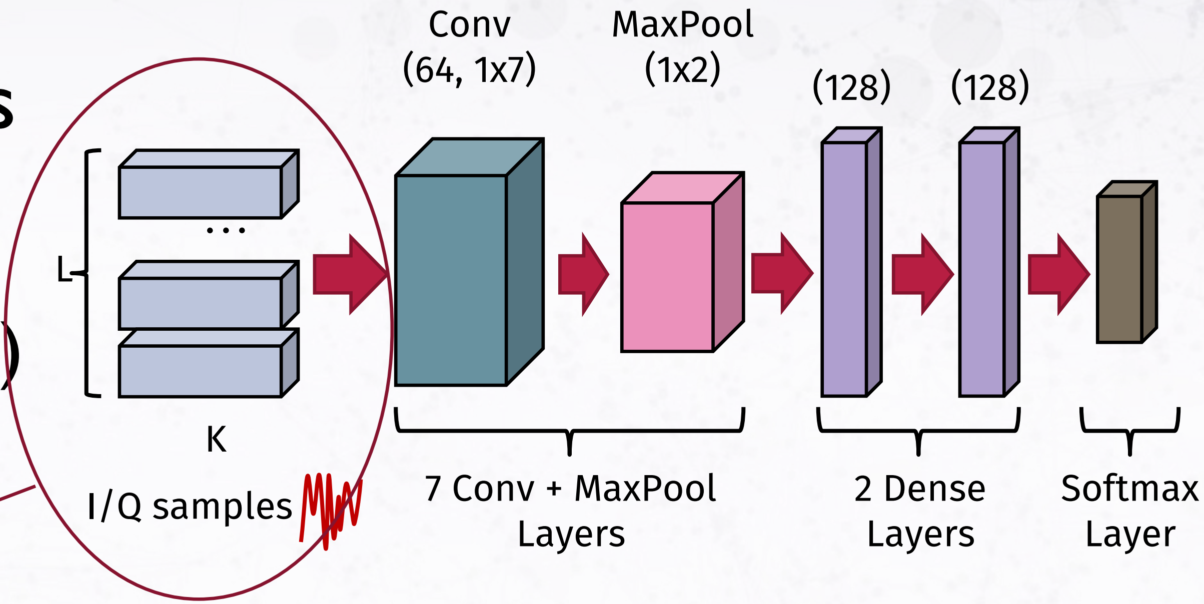
dApps – a possible architecture



Institute for the Wireless Internet of Things at Northeastern

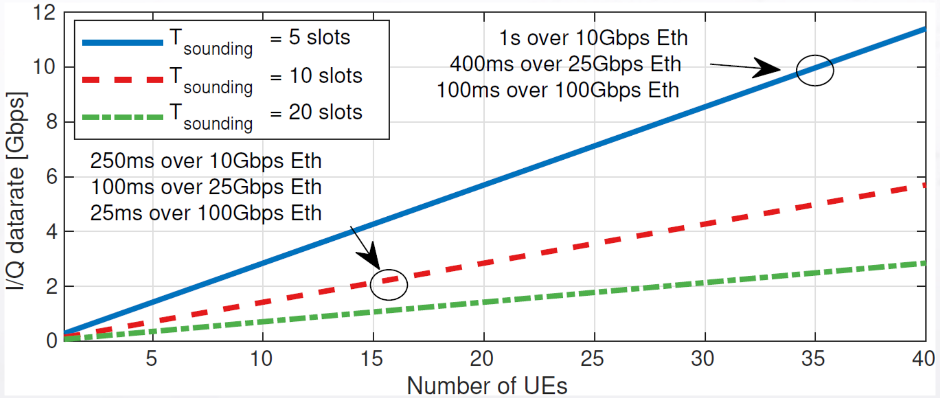
dApps – example of use cases

- I/Q-based deep learning applications
  - Beam management (DeepBeam)
  - Spectrum sensing (DeepSense/Charm)
- Real-time scheduling



Cannot be moved out of the RAN

- security, privacy
- latency



Institute for the Wireless Internet of Things at Northeastern

Resources on NextG open source software

Open 5G Forum – slides and videos online: open5g.info/open-5g-forum
(supported by ACM SIGMOBILE)

Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges

Michele Polese, Leonardo Bonati, Salvatore D'Oro, Stefano Basagni, Tommaso Melodia

Call for papers: IEEE JSAC special issue on Open RAN
https://tinyurl.com/jsac-oran

openrangym.com

Institute for the Wireless Internet of Things at Northeastern University

Intelligent networks with Open RAN

Challenges and opportunities

Michele Polese

Institute for the Wireless Internet of Things

Northeastern University

m.polese@northeastern.edu

with Leonardo Bonati, Salvatore D'Oro, Stefano Basagni, Tommaso Melodia

Partially supported by NSF Grants CNS-1925601, CNS-2120447, and CNS-2112471 and ONR Grant N00014-20-1-2132