Grenze International Journal of Engineering and Technology, January Issue

CCIR: The Next Frontier in Mobile Network Evolution - Integrating Communication and Computing for Enhanced Services

E. Babu Raj[1], Padmalal S[2], Ram Mohan N. R[3] and B. Shaji[4]
[1]Associate Professor, Department of Computer Science and Engineering, CHRIST University, Bengaluru
Email: baburaj.e@christuniversity.ac.in
[2]Professor, Department of Computer Science and Engineering, Mangalam College of Engineering, Ettumanoor, Kottayam
Email: Dist, Kerala, India.
[3]Professor, Department of Computer Science and Engineering, Royal College of Engineering and Technology,
Email: Akkaikavu, Thrissur, Kerala, India.
[4]Associate Professor, Department of Computer Science and Engineering, Nehru College of Engineering and Research Centre, Pampady, Thrissur, Kerala, India.
Email: shajibensiker@gmail.com,shaji4176.cse@ncerc.ac.in

***Abstract*—The mobile RAN is the bridge and enabler between end users and application services, and fortunately, the computational capacity of base stations in RANs[1] has experienced tremendous growth accompanied by increasingly stringent service requirements from emerging applications such as AR/VR.**

These two trends imply an unprecedented opportunity that the abundant computing resources in base stations could be leveraged to host latency-sensitive applications if being managed properly, thereby giving rise to a new vision named Communication and Computing Integrated RAN (CCIR), where not only communication but also computing services are delivered by RANs in a coherent way[2].

In fact, CCIR implies an even more radical departure of designing future mobile networks—going beyond regarding the role of RANs as connecting links. This article provides an elaboration on the fundamental design philosophies and principles of CCIR, the logical architecture.

We will explore further on key tech-nologies toward realizing our vision. Specifically, we provide a thorough view on what CCIR really means – it involves different integration granularities between communication and computing functions; meanwhile it keeps evolving until approaching real-time joint scheduling and holistic resource management based on one unified infrastructure.

To assess the feasibility and advantages brought by CCIR, several field experiments were carried out where computing resources are migrated within different base stations. In conclusion, CCIR is a prospective evolution direction for future mobile networks dealing with communication-computation-converged requirements in new service use scenarios[3].

Using current infrastructure more intelligently and efficiently will facilitate better user experience, greener mobile networks as well as serve as better platform for future advancements in wireless technology.

Grenze ID: 01.GIJET.11.1.343

I. INTRODUCTION

The new emerging services like cloud gaming and metaverse put forward new challenge for 5G and future mobile network that is not only communication oriented but computationally also. For such purpose, there have been some novel RAN technologies proposed to support the diverse scenarios that include URLLC and mMTC[4]. As for computational aspect, MEC has provided the possibility to deploy application services at the edge of network so as to enhance end users QoE. However, with traditional view, the computational infrastructure of both MEC and RAN are separated. Academic research community has expressed their interest in the convergence of communication and computing by proposing new edge computing infrastructure for achieving better latency and resource optimization. One such noticeable approach is the introduction of radio computing node (RCN) and a computing plane that performs computational functions along with communication To this end, this article introduces the concept of CCIR in RANs evolution[5]. We present the design goals and propose an architectural blueprint for realizing CCIR possibilities focusing on deep convergence by designing underlay adoption which ensures control on computing operations via air-interface signaling. This work explores potential key technological components to achieve our proposed architecture for CCIR through an experimental field trial study demonstrating AI task scheduling feasibility through two coordinated base stations followed by open-issue discussion related to standardization activities. The article expands the introduction by highlighting how important is the user equipment participation in integrated communication and computing [6], the continuity and mobility of computing services, as well as the efficient use of RAN's computing resources to transform RAN into a platform that provides not only connectivity but also advanced computing and AI services. This motivates an in-depth study on how CCIR can play this role in helping the evolution of mobile networks.

II. A REVIEW ON LATEST PARADIGM SHIFT FOR MOBILE NETWORK

One notable direction in the evolution of RAN infrastructure is migrating from dedicated hardware towards generic commodity-off-the-shelf (COTS) platforms[7]. .[figure 1] This movement comes up because it was observed that an important part of communication processing is similar to processing required for application services that are hosted on generic servers in centralized data centres as well as edge data centres: such a similarity formed the basis for integrated communications-and-computing-services. In this context, the paper introduces a concept Communication and Computing Integrated RAN (CCIR) which is a RAN that provides both traditional RAN services including connectivity control and mobility control and application services by integrating communication and computing resources. The target of CCIR is to provide more than just connectivity services and to generate new value propositions for network operators. It should be noted that CCIR is not the same as Mobile Edge Computing[8] (MEC)10 1,2,3 which mainly deploys upper-layer application services on COTS servers but does not provide RAN service itself. However, MEC can be deployed on the same infrastructure with RAN tightly interacting via well-defined interfaces too which are corresponding to second level of CCIR integration. The paper focuses mainly on the final form of CCIR (i.e. integration level three) since this is a necessary condition to enable fine-grained scheduling of communication and computing resources, what will increase system performance, improve user experience and increases resource efficiency. Additionally at this level, computing resources and capabilities of user equipment (UE) are fully integrated with CCIR network.

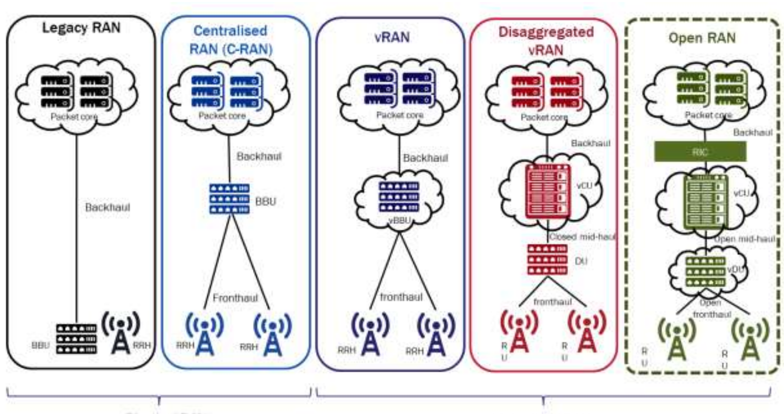

Figure 1: Evalution of RAN

The CCIR vision is a next, natural step in the wireless evolution path towards integrating communication and computing services in the RAN targeting at provision of holistic services. This integration postulate follows from observation, that today's BSs are already exploited ICT resources, which can be leveraged for services' improvement.
One of the main advantages of CCIR is possibility to provide diverse functional use-cases ranging from home-users with their private Virtual assistants to needs of industrial mMTC applications including low-latency processing. By enabling edge-computing in BSs CCIR solutions limit necessity to heavily rely on remote servers what – besides shorter response times- improves privacy protection and security.

III. CCIR – USE CASES

The use cases of CCIR demonstrate how it can revolutionize a mobile network by integrating communication and computing resources. Some of the important ones are mentioned next.
Home Users: New services, requiring low latency as well as high speed processing, could be supported by CCIR femtocells. For example, [9] CCIR femtocells can support personal cloud where the data of the people in the same family are collected and processed locally to build a personalized virtual assistant that manages schedule and gives recommendations with reduced reliance in remote servers for privacy/security reasons.
Vertical Industries: CCIR can easily fulfill the stringent requirements such as ultra-low latency, thus making it a perfect choice for verticals like manufacturing and transportation. By leveraging the existing IT infrastructure, CCIR reduces the expenditure while provides location oriented integrated communication and computation service.
Collaboration Across Devices and Networks: CCIR allows collaboration among mobile devices, RAN and cloud services. It helps in achieving better resource utilization and improved performance by enabling dynamic coordination of computing tasks among various network elements.
Enhanced User Experience: By interconnecting communication with computing, CCIR provides end-to-end service quality in a transparent way. It optimizes resource allocation dynamically so as to satisfy the efficient requirements from users and network.
Sustainability and Efficiency: CCIR exploits the temporal spare computing capacity in RAN to provide computing services, which reduces the need of infrastructure expansion, improves energy utilization efficiency, and processes data near source that also saves energy consumption.
CCIR can be used in multiple areas as evident from the following use cases, where it has shown to improve service, performance and application to different domain.

IV. THE COMMUNICATION AND COMPUTING INTEGRATED RAN (CCIR)- DESIGN PRINCIPLES

The CCIR (Communication and Computing Integrated RAN) architecture, as shown in Fig. 2, is conceptually designed to integrate communication and computing functions inside the Radio Access Network (RAN), so as to create a new network paradigm of mobile networking. This integration operation architecture is basically built on top of the following essential design philosophies:
Underlay Design for Integration: The computational activity in both the base station and User Equipment (UE) can be controlled or driven through air interface signaling[10].
Computing Control Plane and User Plane: This can be achieved by RAN protocols where separate computing control and user planes are established[11].
Heterogeneous Computing Infrastructure Management: The architecture has to manage and schedule efficiently a heterogenous set of computing resources including CPUs, GPUs, NPUs.

V. LOGICAL REFERENCE ARCHITECTURE

The logical reference architecture[12] of the CCIR is a flexible multi layer conceptual framework that describes how the CCIR system is structured and organized. It consists of three main layers:
Network and Computing Infrastructure Layer: This layer constitutes the physical foundation of the CCIR. It includes all hardware and software resources that are used for communication as well as computing. Particularly, it comprises mobile devices with device edge cloud capabilities, RAN nodes with advanced edge computing capabilities, as well as edge/cloud data centers.
Communication and Computing Function Layer: This layer is responsible for implementing the designed communication and computing functions in the RAN, e.g., the dynamic deployment of data offload anchor points to support fast computing-oriented data transmissions among different computing nodes.

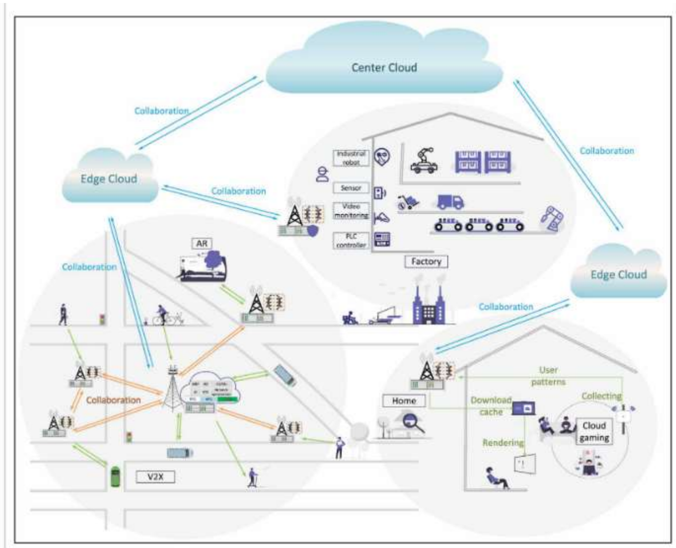

FIGURE 2. Usage scenarios of communication and computing integrated RAN.

Figure 2

VI. Devices with RAN Nodes

The mobile devices are allowed to utilize the computational resource of RAN nodes[13] to conduct AI based tasks as follows.
Offload of computation-intensive task: Mobile devices can offload the computation-intensive tasks such as AI training, AI inference with large models, and video rendering to RAN nodes with integrated computing capabilities.
Collaboration Across Computing Resources: The integration facilitates the collaboration between computing capabilities of mobile devices, RAN nodes, edge cloud and central cloud.
Dynamic Workload Distribution: The RAN can dynamically distribute computing workloads across multi-tier computing locations including mobile devices, RAN nodes, edge clouds and central clouds
Joint Communication and Computing Control: The integration enables powerful joint communication and computing control, scheduling, management, and orchestration mechanisms.
Heterogeneous Computing Infrastructure: RAN nodes are evolving to incorporate embedded computing capabilities e.g., extended computing cards with GPUs, CPUs, NPUs, and DPUs
Radio Cloud Infrastructure: The radio cloud infrastructure for hosting of RAN functions and edge applications, both running concurrently to improve integration and resource utilization efficiency.

VII. Key Technologies

The pursuit of faster speeds, lower latencies, and higher reliability has been the constant in the evolution of mobile networks. The Communication and Computing Integrated RAN (CCIR), is a communication and computing integrated architecture that will be realized in the next-generation mobile network. It is not an incremental improvement as perceived in conventional discussions; rather, it’s a disruptive innovation and new way to organize Mobile Network Operating System.
At the lowest infrastructure layer, some hardware and software components achieve communication and computing functionalities which consists of mobile end user devices with device edge cloud capabilities, RAN nodes equipped with deep edge computing capabilities, edge cloud instances allocated on demand at random network locations along with central cloud.
The network function layer is the engine room of CCIR, and it is where communication and computing functions are integrated. CEF and CCF [14] are the major functions in this layer, where they execute computing tasks and manage the lifecycle of both computing resources and communication resources.
The management orchestration layer manages orchestration for CCIR services. It also manages RAN – Computing resource, RAN Communication & Computing function as well as seamless integrated Service. This

layer exposes service/function/dataset for external consumer through intent driven interface with 3 types of responsibility including service management & exposure function, data management & exposure function and AI/ML management.

VIII. Revolutionizing RAN with Integrated Communication and Computing: Functionality and Interface Design

The evolution of mobile networks has been a journey of continuous innovation towards faster speeds, lower latencies and higher reliability. The Communication and Computing Integrated RAN (CCIR) vision makes the convergence between communication and computing as the new driver for future innovation in 6G. CCIR is not just another possible network architecture but rather a target concept that changes the way we envision what RAN is.

In its core, CCIR emphasizes the redesigning of RAN functionality[15] and interfaces to enable the integration between communication and computing [see figure 3]. A new plane, referred to as computing plane, is introduced to encapsulate functionality that relates to processing of data traffic and managing computing resources within the radio access network (RAN).

The evolution of mobile networks has come to a turning point with the emergence of Communication and Computing Integrated RAN (CCIR) architecture, which not only brings the potential of higher mobile network performance, but changes how communication and computing services are orchestrated and exposed for users/applications. At the heart of CCIR lies on joint orchestration of communication and computing services that is a complex process due to dealing with heterogeneity in terms of resources, functions, and services. This orchestration is realized by Joint Communication and Computing Orchestration (CCO) as a key enabler function in the management and orchestration layer.

IX. Field Trials Conducted

The mobile network evolution has come to a turning point with the Communication and Computing Integrated RAN (CCIR) architecture, which not only brings prospective capacity improvement of mobile networks but also a new way in orchestrating and exposing communication and computing services to users and applications.

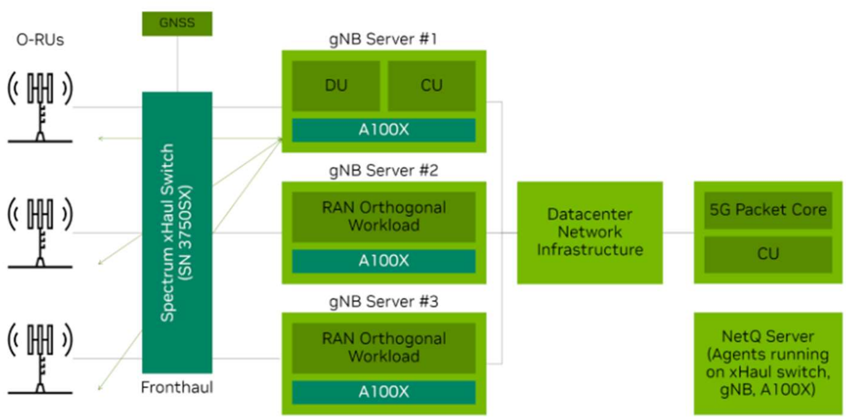

Figure 3: Collocating RAN and AI in the cloud

The Need for RAN Computing Resource Orchestration

With the network becoming more complex and applications that are more [16]data hungry, there is a growing need to efficiently orchestrate the RAN computing resources required to optimize wireless resources and 5G base station performance. However, the current network architecture and 5G base station software-hardware systems does not expose general-purpose computing resources for advanced functionalities such as data analytics and AI computations.

To investigate this challenge, we have performed field trials for validating RAN computing resource orchestration for intelligent RAN. With the help of intelligent Adaptive Modulation and Coding (AMC), the RAN scheduler can aggregate distributed computing resources of neighboring multiple base stations together. The intelligent AMC AI/ML[17] model is a method which intelligently selects modulation and coding schemes based on varying channel condition, but it needs to be trained or updated periodically because radio link conditions are time varying. Finally, real-time optimized MCS decisioning via the trained models will improve both UE QoE and network spectrum efficiency.

Three base stations in total formed a RAN computing cluster, where one acted as master node to control computing

The field trials on RAN computing resource orchestration for intelligent RAN demonstrated that the CCIR is able to execute AI-related tasks over tightly-interworked radio network of base stations towards autonomous networks in which communication and computing are so deeply interwoven in the RAN that a truly pervasive service experience beyond connectivity can be made available. For industrial standardization and commercial deployment, the promising solutions proposed by different contenders are likely to become more mature as they all strive at realizing the full potential of CCIR.

RAN Node			Before AI		After AI migration	
			CPU Load	Memory	CPU Load	Memory
gNB1	ID 11581806	Total	25.2%	7431M	13.07%	7177M
		Intelligent AMC App	12.13%	254M	0	0
gNB2	ID 11710921	Total	4.07%	7873M	15.65%	8153M
		Intelligent AMC App	0	0	11.58%	280M

Table 1: Computing the resource usage

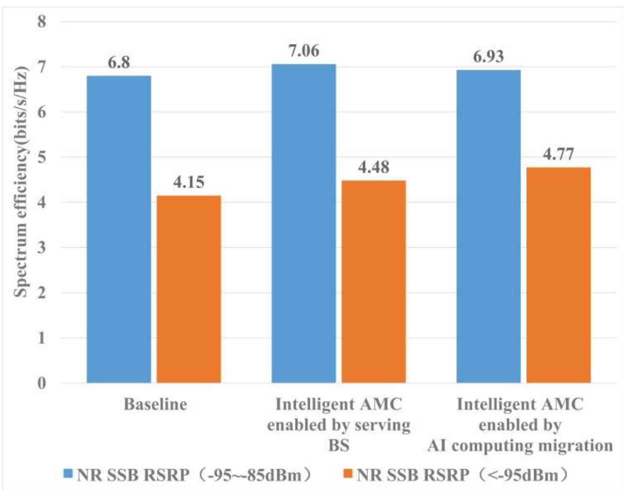

Figure 4: Spectrum Efficiency Comparison

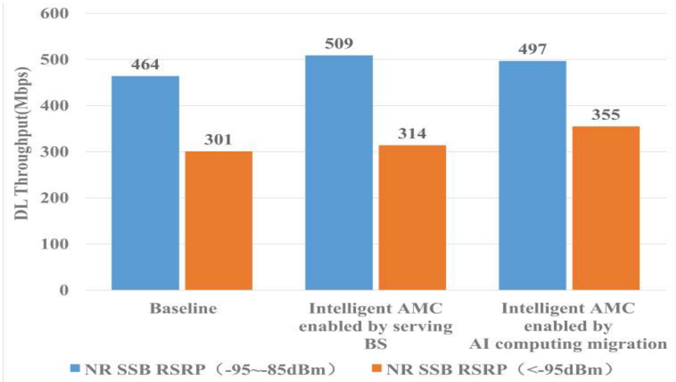

Figure 5: Throughput Comparison

X. CONCLUSION

The evolution of mobile networks has entered into a new era with the proposition of Communication and Computing Integrated RAN (CCIR) architecture . It is an architectural disruptive innovation that co-integrates communication and computing in the RAN to satisfy emerging ultra-reliable low-latency communication (URLLC) requirements. CCIR means not only faster internet, but also intelligence: RAN that can think and act in real-time. AI service is abstracted into executable tasks, such as data acquisition, model training, and distributed inference, with end-to-end AI service guaranteeing orchestrated execution of AI task and user-plane functions. The integration of AI/communication is one of three primary usage scenarios outlined in the requirement document for IMT-2030, where CCIR leads enabling RAN-native intelligent services.

The Management and Orchestration Layer manages/orchestrates RAN computing resources, communication function, computing function, and integrated service. It provides service exposure and allows RAN services to be discovered by external consumers through intent-driven interfaces so that any perceived RAN-servicelogic is totally hidden from the consumer point-of-view making it as simple as possible (onboarding/off-boarding) for the RAN service consumer.

Based on field trial result, we observed that CCIR could realize automation of network operation/resource optimization with intra-node AI task of properly synchronized B/Ss. This result implicates promisingness of CCIR's capabilities regarding network operation Automation/Resource Optimization. Nevertheless, standardization of CCIR is quite challenging and needs contributions from wider stakeholders to ensure cross-industrial adoption/implementation globally.

As we're heading to the 6G era, CCIR is an innovation and new mobile network conception in itself. From technology perspective, CCIR is a must-evolutionary-step of RAN going forward for physically integration of communication and computing, which has never been no longer a vision but existing. But how to standardize? How should it work? How far it can be? All those questions are on people side.

REFERENCES

[1] Checko, Aleksandra, et al. "Cloud RAN for mobile networks—A technology overview." IEEE Communications surveys & tutorials 17.1 (2014): 405-426.

[2] Spinuzzi, Marc A., and A. Marc. "CCIR for complex and uncertain environments." Unpublished monograph. Ft. Leavenworth, KS: School of Advanced Military Studies 1 (2007).

[3] Huang, Yuhong, et al. "Communication and Computing Integrated RAN: A New Paradigm Shift for Mobile Network." IEEE Network (2024).

[4] Pokhrel, Shiva Raj, et al. "Towards enabling critical mMTC: A review of URLLC within mMTC." IEEE access 8 (2020): 131796-131813.

[5] Li, Rongpeng, et al. "TACT: A transfer actor-critic learning framework for energy saving in cellular radio access networks." IEEE transactions on wireless communications 13.4 (2014): 2000-2011.

[6] Von Eicken, Thorsten, et al. "Active messages: a mechanism for integrated communication and computation." ACM SIGARCH Computer Architecture News 20.2 (1992): 256-266.

[7] Basili, Victor R., and Barry Boehm. "COTS-based systems top 10 list." Computer 34.5 (2001): 91-95.

[8] [8] Mao, Yuyi, et al. "A survey on mobile edge computing: The communication perspective." IEEE communications surveys & tutorials 19.4 (2017): 2322-2358.

[9] Abbas, Nasir, et al. "Mobile edge computing: A survey." IEEE Internet of Things Journal 5.1 (2017): 450-465.

[10] Alam, Tanjir, Sreenath Reddy Thummaluru, and Raghvendra Kumar Chaudhary. "Improved multifunctional MIMO cognitive radio system for integrated interweave-underlay operations." IEEE Transactions on Microwave Theory and Techniques 70.1 (2021): 631-640.

[11] Xie, Junjie, et al. "Control plane of software defined networks: A survey." Computer communications 67 (2015): 1-10.

[12] Boella, Guido, and Leendert van der Torre. "A logical architecture of a normative system." International Workshop on Deontic Logic and Artificial Normative Systems. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.

[13] Habibi, Mohammad Asif, et al. "A comprehensive survey of RAN architectures toward 5G mobile communication system." Ieee Access 7 (2019): 70371-70421.

[14] Li, Xiang-yang, et al. "Laparoscopic management of cholecystoenteric fistula: a single-center experience." Journal of International Medical Research 45.3 (2017): 1090-1097.

[15] Ionescu, Ana-Cristina. "How can Corporate Social Responsibility (CSR) and Innovation enhance enterprise performance?: possible pathways for the Romanian SMEs in the context of the global economic crisis." Cases on SMEs and Open Innovation: Applications and Investigations. IGI Global, 2012. 163-197.

[16] Luo, Hao, and Hung-Yu Wei. "Resource orchestration at the edge: Intelligent management of mmwave ran and gaming application qoe enhancement." IEEE Transactions on Network and Service Management 20.1 (2022): 385-399.

[17] Rial, Alvaro Valcarce, et al. "The Role of AI in 6G MAC." Authorea Preprints (2023).