Towards AI Powered and Secure Carrier-Grade Open RAN Platform

Future RAN Competition Project

Final project report

Executive summary

The telecommunications sector is undergoing a significant transformation with the development of the "Towards AI Powered and Secure Carrier-Grade Open RAN Platform" project. This collaboration between Microsoft UK, Intel R&D UK Ltd., Capgemini, and The University of Edinburgh, funded by the UK government's DSIT, aims to revolutionize mobile networks through Open RAN, integrating AI and cloud security into a flexible, carrier-grade solution.

The project has successfully implemented a scalable 5G network and developed a far-edge RAN platform with real-time Linux kernel optimizations and Kubernetes abstractions, enabling quick provisioning and vendor diversity. REST APIs have been instrumental in simplifying network deployments, allowing a functional platform to be set up in just 30 minutes and facilitating the installation of RAN CNF as marketplace apps.

Energy efficiency and network performance have been enhanced through applications that manage power consumption and interference. AI-driven analytics contribute to improved network troubleshooting and anomaly detection. Security measures, including a ransomware attack detection app, align with the best DevSecOps practices, ensuring robust defense mechanisms.

A pioneering aspect of the project is the dynamic service models using eBPF, which allow for real-time, flexible telemetry collection, showcased by the Janus prototype. This innovation promises greater flexibility for future Open RAN development.

The project's broader impact is evidenced by its recognition at industry events and contributions to academic research. Despite supply chain challenges, the team's commitment to diversified sourcing and influencing standards illustrates a strategic approach to shaping the future of Open RAN. In line with the UK's diversification strategy, the project promotes a competitive vendor ecosystem and fosters local talent development. It represents a forward-thinking step in creating a more open, efficient, and innovative telecommunications infrastructure.

Introduction

In the fast-paced world of telecommunications, the race is on to harness emerging technologies that promise to revolutionize the way we connect to the internet. Topping the list of these innovations is Open Radio Access Network (Open RAN), a groundbreaking way of building mobile networks, set to redefine the rules of the game.

Traditional mobile networks have relied on closed hardware and software provided by a select few vendors. These systems were tightly integrated, making upgrades or modifications a costly and cumbersome affair. Open RAN came in as a game-changer, separating the software from the hardware, digital from analogue, thus making mobile networks more adaptable, cost-effective, and open to a broader spectrum of vendors.

Now, think of infusing this Open RAN with the power of Artificial Intelligence (AI). The ability of AI to learn from data and make intelligent decisions could significantly enhance the performance

and management of Open RAN. It could detect anomalies, optimize resource allocation, reduce power consumption, and even predict and prevent potential problems before they occur. Imagine securing this AI-powered Open RAN with the robustness of the cloud. This was precisely the vision that led to the inception of the project titled "Towards AI Powered and Secure Carrier-Grade Open RAN Platform".

The project was born out of the collaboration between four industry and academia leaders: Microsoft UK, Intel R&D UK Ltd., Capgemini, and The University of Edinburgh. Each partner brought to the table their unique expertise in cloud technology, AI, cybersecurity, academic research, and telecom infrastructure. The project is co-funded by the UK government DSIT (Department for Science, Innovation and Technology) Future RAN diversification funds. Together, we envisioned a carrier-grade cloud solution that would empower operators to deploy Open RAN network functions easily and securely.

Project goals and achievements

The project set its sights on five key areas: manageability, observability, automation, efficiency, and security. Each of these areas came with its own unique set of challenges and questions.

Building a 5G network is no mean feat. To simplify the process, the project developed hardware and software blueprints that were tailored to meet the unique challenges of building a 5G network, such as successfully connecting and integrating various components (servers, switches, clock sources, radio) together, provisioning and life-cycle management, as well as ensuring consistent connectivity and high-speed connections of many mobile devices.

We deployed a building-wide 5G enterprise network over 5 floors with remotely tens of controllable mobile devices spread around the building. We implemented a remote access to the lab that allowed collaborators from different locations to run experiment on the network, and jointly work on the problems.

One of the challenges that the project tackled head-on was the development of a far-edge RAN platform that could meet various demands of virtualized RAN. This platform needed to support r eal-time Linux kernel, and to be optimized to minimize latency and support the various low-laten cy requirements put forth by different vendors. To achieve this, the project team had to dive deep into the complexities of platform performance and develop strategies to optimize every aspect of its operation.

The platform also had to support Kubernetes and provide Kubernetes abstractions that are requir ed by different vendors. These abstractions include isolated CPUs, huge pages, and virtual functio ns, among others. We spent time understanding how best to express these abstractions and redu ce the number of errors that can occur in this provisioning.

To provision and configure the platform in a flexible way, the team developed a system that uses REST APIs for remote provisioning. This system allows for the provisioning of bare metal to a func tional far-edge platform in a mere 30 minutes. This is a significant achievement, as it allows netw orks to be deployed much more swiftly, efficiently and at scale, than traditional methods.

Alongside the platform provisioning, the team also addressed the challenge of installing RAN CN F (Cloud-Native Functions) as an app from a marketplace. Achieving this required building anoth er REST API for RAN provisioning and testing it with a full containerized implementation of Capge

mini's 5G RAN deployed through the Azure Arc platform. We further tested the platform API with several other vendors' RAN CNF. The team also had to create advanced platform features for sec urity and correctness, ensuring a seamless and verifiable deployment of RAN CNF.

All these elements combined to create a far-edge RAN platform that is not only robust and efficie nt but also flexible and adaptable, able to meet the diverse needs of various vendors. This platfor m forms the backbone of the project's vision for a next-generation telecom network powered by AI and secured by the cloud.

Increasing energy efficiency was another challenge that the project tackled head-on. We developed applications that not only reduced RAN power consumption, even during peak hours, but also detected external interference and avoided inter-cell interference. These innovations led to more efficient use of network resources and a significant improvement in network performance.

Troubleshooting and system visibility were other key areas the project focused on. Advanced algorithms, driven by AI and machine learning, were used to analyse the vast amounts of data generated by the network, detect anomalies, and provide insights for effective troubleshooting. In the prototype we built, we were able to collect hundreds of telemetry sources from RAN and platform several times per second. We fed this data to our custom-built generative AI model that can process the data in real-time and detect any anomalies and their root causes.

Security is always a significant concern in any network, both in terms of software development and deployment. To address these, we leveraged the best DevSecOps industry practices, including pipelines, unit and system tests, scans, and verifications. Furthermore, we developed a ransomware attack detection app that can swiftly detect a ransomware attack at a telco far-edge, and repair compromised servers.

One of the highlights – dynamic service models

applications we discussed, is the idea of dynamic service models. These models provided a workaround to the restrictions of current xApps that operate on the non-real-time Radio Intelligent Controller (nRT-RIC) and are limited by static E2 service models. These static models are defined by industry standards and any alteration to them necessitates an extensive standardization process.

The dynamic service model we developed, on the other hand, offer a more flexible solution. They allow xApp developers to utilize Extended Berkeley Packet Filter (eBPF), a Linux-based technology, to create their own methods for collecting internal RAN telemetry. A developer can express their service model through a custom code compiled with the eBPF framework, which is both secure and verifiable, and can be seamlessly deployed inside a RAN at run-time.

To demonstrate this, our team built a prototype called Janus. Integrated with Capgemini and Intel RAN components, Janus showcased the potential of dynamic service models and served as the foundation for several applications. This innovative approach brings a new level of flexibility to service models, opening up a realm of possibilities for the future of Open RAN.

Broader impact and lessons learned

The project's impact was not confined to the lab. It made its presence felt at trade shows like MWC Barcelona, Open RAN world events and IntelON, demonstrating to the wider industry the

achievements of the project. The team received recognition in the form of the Light Reading Editor's choice award for AI/ML in RAN in 2023, and we were also finalists in the Fierce Telecom award for AI/Analytics/Automation in the same year.

In terms of academic contributions, the team published three papers in the prestigious ACM Mobicom conference, with one more under review. We also published a white paper aimed for the industry decision-makers, contributing significantly to the body of knowledge on Open RAN. This material is available on the project web site.

During the project, we encountered several challenges, primarily on the supply chain side. Some components experienced long delays of more than 6 months. It was often difficult to purchase Raspberry Pi devices at scale, which we used as mobile clients for our building-wide deployments. Yet, these hurdles only served to underscore the importance of having a diverse supply chain and drove the team to seek out alternative solutions.

In shaping the future of Open RAN, the project also recognized the importance of influencing standards and policies. It ran a dissemination campaign to promote the benefits of the platform and analytics and interacted with ecosystem partners to gain understanding how to shape the project outcomes into standards which aligns with DSIT Open RAN principles.

The project's alignment with the UK DSIT diversification strategy was another achievement worth mentioning. By building the platform with open interfaces, it allows a larger pool of vendors to participate. This not only fosters competition but also drives innovation. The project also has an educational aspect, as its collaboration with a major UK university helps develop UK talent in the area of Open RAN, further supporting the DSIT’s diversification strategy.

In summary, the "Towards AI Powered and Secure Carrier-Grade Open RAN Platform" project is a significant step forward in the development of Open RAN technology. It brings together the power of AI, the security of the cloud, and the flexibility of Open RAN to create a next-generation telecom network. With its significant technical achievements, and support for the UK DSIT Diversification Strategy, the project hopes to leave a lasting impact on the world of telecommunications.