ANTS 2024
15–18 December 2024, Assam, India

Resource Profiling for Virtualized Radio Access Networks

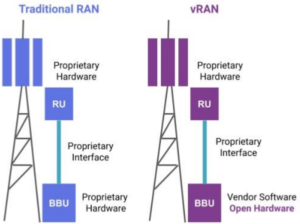
Rajashekhar Reddy Tella*, Utkarsh Gupta*, Gudepu Venkateswarlu*, Venkatarami Reddy Chintapalli†, Bheemarjuna Reddy Tamma‡, Koteswararao Kondepu*

* IIT Dharwad, Karnataka, India
† NIT Calicut, Kerala, India
‡ IIT Hyderabad, Telangana, India

Outline

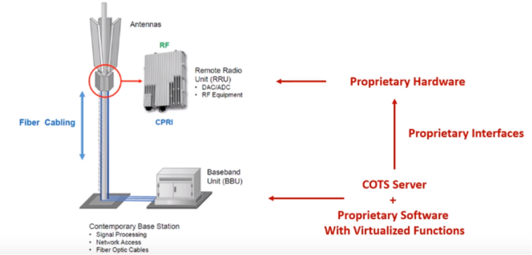
Introduction

Introduction to vRAN



- The transition to vRAN from traditional RAN
- Replacing specialized hardware with **Commercial Off-The-Shelf (COTS)** hardware powered by **general-purpose processors.**
- Adoption of Network Functions Virtualization (NFV) enables: Ease of deployment, Increased flexibility and Cost savings.

Significance of vRAN



1. Deploys RAN functions as Containerized Network Functions (CNFs).
2. Supports 5G and beyond with higher efficiency and scalability compared to traditional RAN.
3. Multiplexes multiple base station workloads on shared hardware, allowing co-location with other workloads.

Challenges in vRAN Deployment

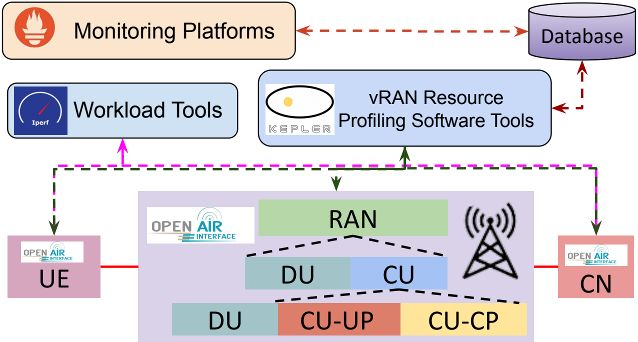
Consolidated environments lead to resource contention for:

1. CPU cores
2. Last Level Cache (LLC)
3. Memory Bandwidth

Ensuring consistent performance of vRAN workloads without compromising Quality of Service (QoS) for an ever-growing number of subscribers, in this consolidated environment is challenging due to the high variability and unpredictability arised due to contention of resources.

Research Objectives

- Understand how next-generation scalable processors can optimize vRAN components for all scenarios of RAN.
- Analyze the effect of controllable resource knobs on performance of RAN, including:
  - CPU pinning
  - LLC allocation
  - Memory Bandwidth
  - Energy consumption



UE - User Equipment; RAN - Radio Access Network; DU - Distributed Unit; CU - Centralized Unit; CU-UP - CU User Plane; CU-CP - CU Control Plane; CN - Core Network

Figure: System Architecture used in carrying experimentation

Control Knob: CPU Core

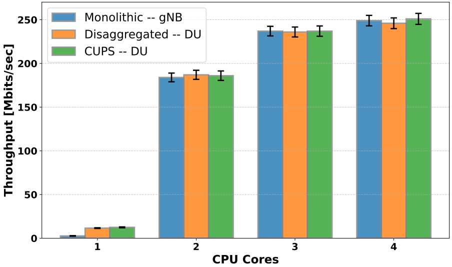


Figure: Throughput vs CPU Cores

Observation:

- The trend is generally upward, but at least 1 CPU core is needed to avoid low throughput.
- Throughput improves with more cores, but beyond 3 cores, the increase is minimal, indicating diminishing returns.

Control Knob: Memory Bandwidth

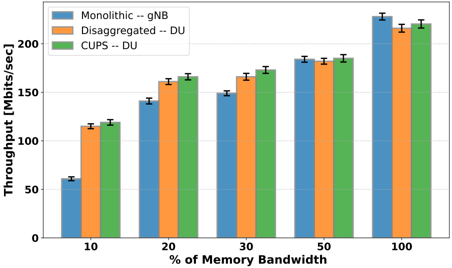


Figure: Throughput vs Memory Bandwidth

Observation:

- The trend is generally upward
- Throughput at 10% memory bandwidth is half the throughput at the 100% memory bandwidth, whereas in monolithic, it is fourth.

Control Knob: Last Level Cache ways

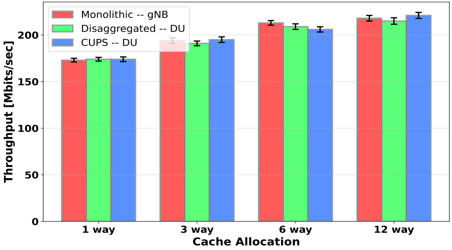


Figure: Throughput vs Cache Allocation

Observation:

- The throughput increases with the increase in cache ways allocated to the RAN pods
- If there is a high demand for cache, one can use a higher number of CPUs at a lesser cache way to achieve the same performance.

Power Considerations

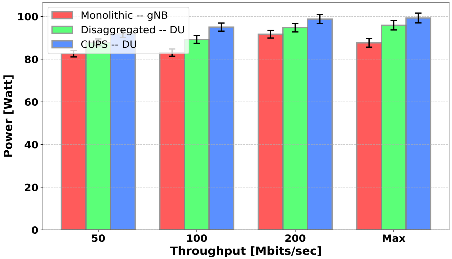


Figure: Power vs Throughput workload

Observation:

- Due to efficiency of the hardware used, there isn't much variation observed in power consumed.
- CUPs scenario consumes more power compared to the remaining two for a givenworkload.

Power Heatmap

Power in Watt												
CPUs Allocated	2 CPUs				3 CPUs				4 CPUs			
Cache Ways	1	6	9	12	1	6	9	12	1	6	9	12
Throughput [Mbits/sec]												
10	124	118.14	114.14	103.14	105	97.47	94.04	95.78	86.91	86.85	84.65	80.54
50	119.28	115.28	115	118.28	107.71	100.55	99.37	96.04	93.08	89.32	88.14	81.45
100	109.85	109.21	110.57	112	107	104.35	103.77	99.65	95.55	92.37	91.27	88.88
150		105.57	105.85	108.28	104.28	103.88	105.05	101.78	98.71	94.84	93.62	91.48
200						104.4	107.48	103.18	100.58	97.31	94.8	91.71
230								106.77	NA	97.44	95.05	91.3

Figure: Heatmap for different allocation of CPU and LLC ways (Monolithic Scenario)

Observation: We observe that to achieve the same throughput with minimal power consumption, it’s necessary to use the maximum available resources, such as the highest number of CPU cores and LLC ways. Thus, we see a tradeoff between power and resource usage.

Conclusion

- Profiled the RAN component of cellular networks.
- Results indicate allocation of CPU cores, LLC and Memory have a significant impact on achieving throughput.
- Optimal resource allocation can enhance performance among the considered parameters

Future work:

- Develop power-saving configurations.
- Prioritize power efficiency of the RAN.
- Schedule resources based on usage patterns for optimum use of resources.

Thank You!