Hackathon Problem Statement:

Intelligent Fronthaul Network Optimization: Topology Identification and Link Capacity Estimation

Description:
The fronthaul TSN network (as defined in O-RAN) plays a critical role in modern mobile networks, connecting Distributed Units (DUs) and Radio Units (RUs) to ensure seamless communication and efficient resource utilization. With increasing traffic demands and the complexity of network topologies, optimizing the fronthaul network is essential to reduce the total cost of ownership.

Participants are tasked with developing an innovative solution to optimize fronthaul network performance by addressing the following challenges:

Challenges:

1. **Network Topology Identification:**
   - Analyze the fronthaul network traffic patterns based on the provided historical data log to identify the topology, focusing on identifying the cells that share the same links.
   - Develop tools to dynamically visualize and map the fronthaul network topology using the available historical data**.**
2. **Link Capacity Estimation:**
   - Estimate the required Ethernet link capacity for groups of cells identified during topology analysis, considering traffic patterns.
   - Account for traffic capture points at the DU side before the congested link and at the RU side after the congested link to estimate and mitigate congestion.

Here is the link to the Historical Data Logs to be analyzed: https://tinyurl.com/52xbe4n9

Hints for Using Historical Data Logs:

1. Historical data in the throughput file is at the symbol level. Conversion of symbols to slot is done follows:
   - 1 slot = 14 symbols = 500 microseconds.
   - 1 symbol = 500 microseconds / 14 = 35.7 microseconds.

2. Cells sharing the same link will exhibit correlated packet loss during overload at the RU.
3. Timing in the "Packet Statistics" file (taken at the RU side) may be shifted due to transmission delays and clock inaccuracy, compared to sent timing at DU.
4. For estimation of link capacity use file with throughput per symbol and consider available buffer size to detect when packets start get dropped by the switch.
5. More details of the format of the file are provided in the readme file stored in the repository. There are dedicated throughput and Packet Statistics file for each cell, with total 24 cells.

Figure 1 illustrates an example of the correlation of packet loss among a group of cells sharing the same link during periods of traffic congestion. This correlation helps in identifying which cells are connected to the same Ethernet link.

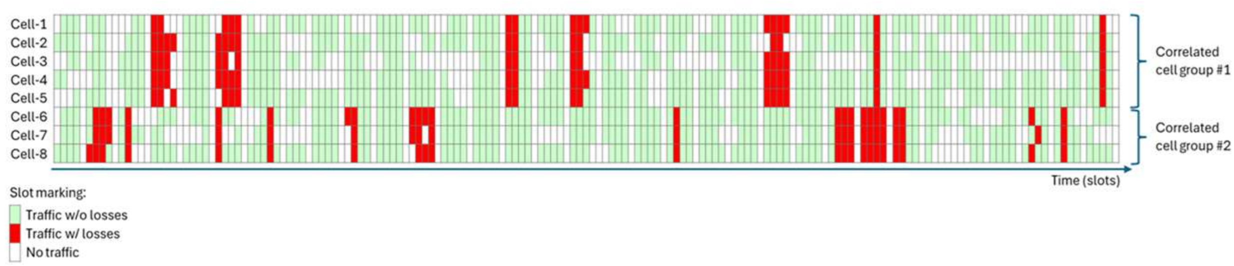

Figure 1: Example of pattern of FH congestion events in time(slots) for each cells.

Fronthaul network topology to be analysed

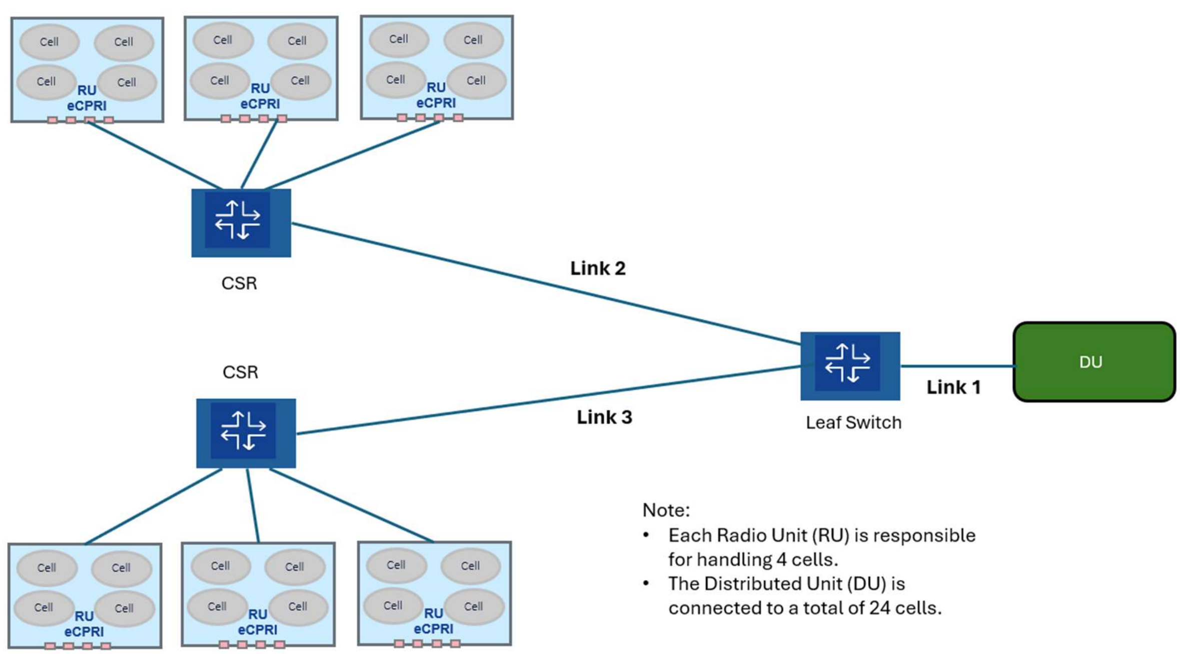

Figure 2: Fronthaul Network

Expected Outcomes:

- Develop a solution that dynamically identifies and visualizes the fronthaul network topology. This solution should map cell numbers to their respective links (Link 1, Link 2, and Link 3 as shown in **Figure 2**) using the provided historical data logs and partial information about the links, such as:
    - Cell1 is connected via Link2. Cell2 is connected via Link3.
- Generate traffic pattern snapshots similar to those shown in **Figure 1** for identifying cells sharing the same link.
- Accurately estimate the total link capacity requirements on Down link for Links 1, 2 and 3, as depicted in Figure 2 for optimum network without congestion, <u>for two cases</u>: with buffer and without buffer.
    - For buffer case: Total buffer size at leaf switch is 4 symbols (i.e.143 microsecond). Note: Buffer size expressed as time interval can be converted to size in bits by multiplying time interval to link rate.
        - Example for 25Gbps link rate, would be 143us x 25Gbps = 3.575Mbits.
    - Packet losses are permitted for up to 1% of the slots carrying traffic per cell
    - **Note:** When estimating optimum link capacity needs consider both buffer size the switches and acceptable packet loss.
- Calculate the data rate (in Gigabits per second, Gbps) per slot for the aggregated traffic of cells for Links 1, 2, and 3. Generate a graph similar to the example shown in **Figure 3**, <u>where the time resolution is one slot over a period of 60 seconds</u>. (Hint: Gnuplot can be used for graph generation.)

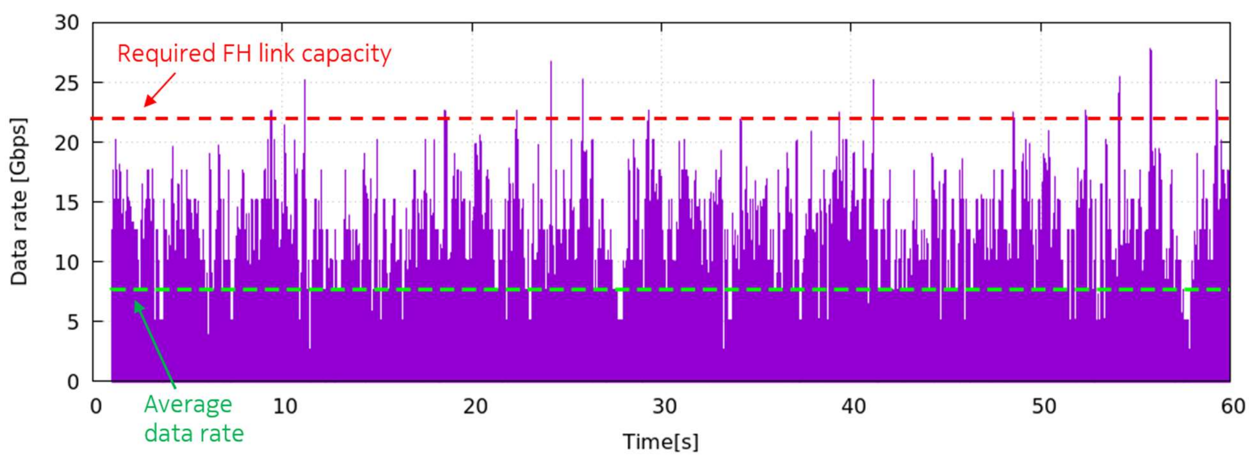

Figure: 3.

Additional information:

Real-World Impact:

- Accurately identifying the required fronthaul bandwidth for links 1, 2, and 3 under worst-case traffic scenarios (i.e., a combination of varying traffic loads across cells simultaneously, along with packet multiplexing gains) enables operators to deploy optimal network capacity.
- This approach prevents over-dimensioning based on peak cell capacity for each individual cell without considering actual cell load at the site, which would otherwise lead to significantly higher network deployment costs.

Judging Criteria:

- **Innovation**: Novelty and creativity of the approach.
- **Technical Feasibility**: Accuracy and reliability of topology identification and link capacity estimation.
- **Scalability**: Ability to adapt the solution to larger or more complex networks.
- **Impact**: Potential to optimize fronthaul network and reduce congestion.
- **Presentation**: Clarity and effectiveness of the solution demonstration.

Tools and Technologies:

- AI/ML Usage: Participants are encouraged to use AI/ML techniques to address the challenges of topology identification and link capacity estimation.
- Dynamic visualization tools can be used for network topology mapping.