Project #4-5 – Packet Loss Event Classification
Background

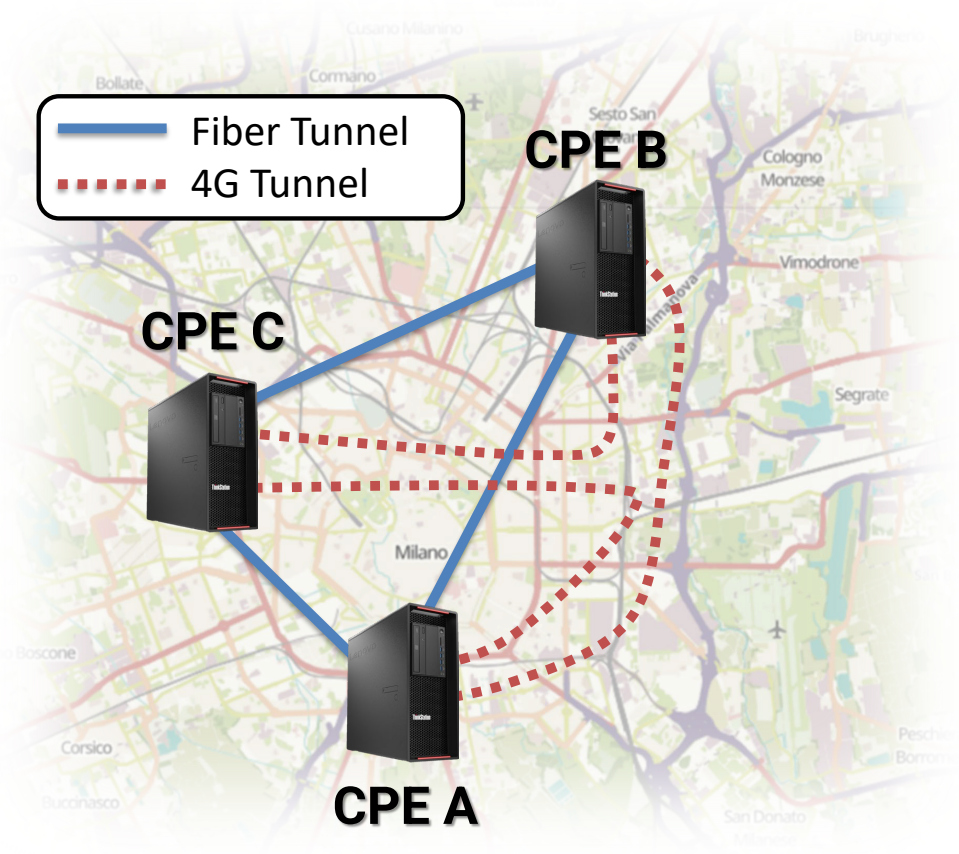
- SD-WAN (Software-Defined Wide Area Network) is an emerging technology for the interconnection of geographically distributed company branches
  - Uses **one or more public best-effort internet access** to interconnect the different sites through one CPE (Customer Premises Equipment) in each site. The interconnection is secure by conveying data through encrypted tunnels.
  - Each of these best-effort internet access corresponds to **different paths in the public Internet**

- **Problem:** Mission-critical inter-site traffic requires strict QoS (e.g., extremely low packet loss). Network inefficiencies (latency spikes, packet loss) on the paths can cause performance degradation and potential financial impact (e.g., in financial contexts).
- We would like to predict network inefficiencies (e.g. latency spikes or packet loss bursts) before they happen so as to redirect traffic on the best path

- Can we use ML to predict packet loss before it happens based on previous historical data?

Project #4-5 – Packet Loss Event Classification
Testbed setup

- Three geographically distributed SD-WAN CPEs across Milan
- Dual WAN connectivity per CPE:
  - Fiber (FTTC/FTTH) tunnel
  - 4G cellular network tunnel
- A latency measurement campaign has been performed on this setup



Project #4-5 – Packet Loss Event Classification
Dataset

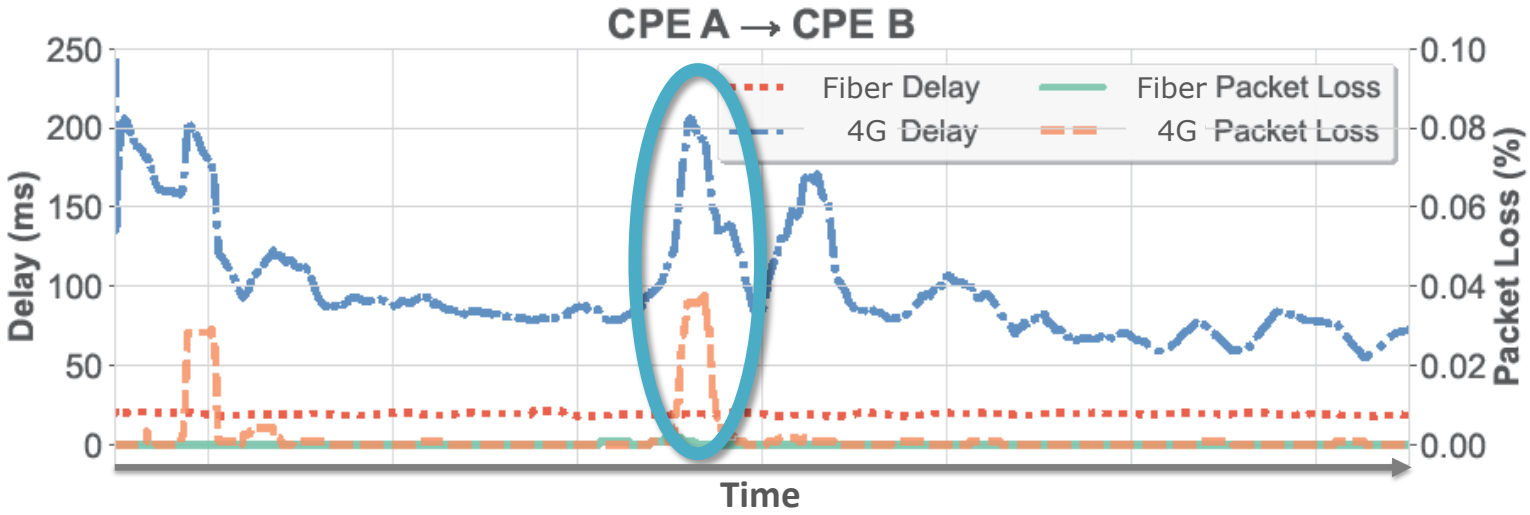
- For any ordered pair (CPE X, CPE Y), two measurement files (.csv) have been generated (one for the fiber tunnel and one for the 4G tunnel)
- Two measurement windows, respectively of duration ~15h and ~ 30h, with one second granularity

	time	delay_ms
23691	2025-03-05 19:05:22.567	13.2708
23692	2025-03-05 19:05:23.567	21.1522
23693	2025-03-05 19:05:24.566	14.8669
23694	2025-03-05 19:05:25.569	-1.0000
23695	2025-03-05 19:05:26.569	14.7499
23696	2025-03-05 19:05:27.570	13.8136
23697	2025-03-05 19:05:28.570	13.6749

Single file structure
(e.g. fiber tunnel from CPE A to CPE B)

Packet losses are identified by the value **-1** in the **delay_ms** column

Project #4-5 – Packet Loss Event Classification
The Dataset Visualized

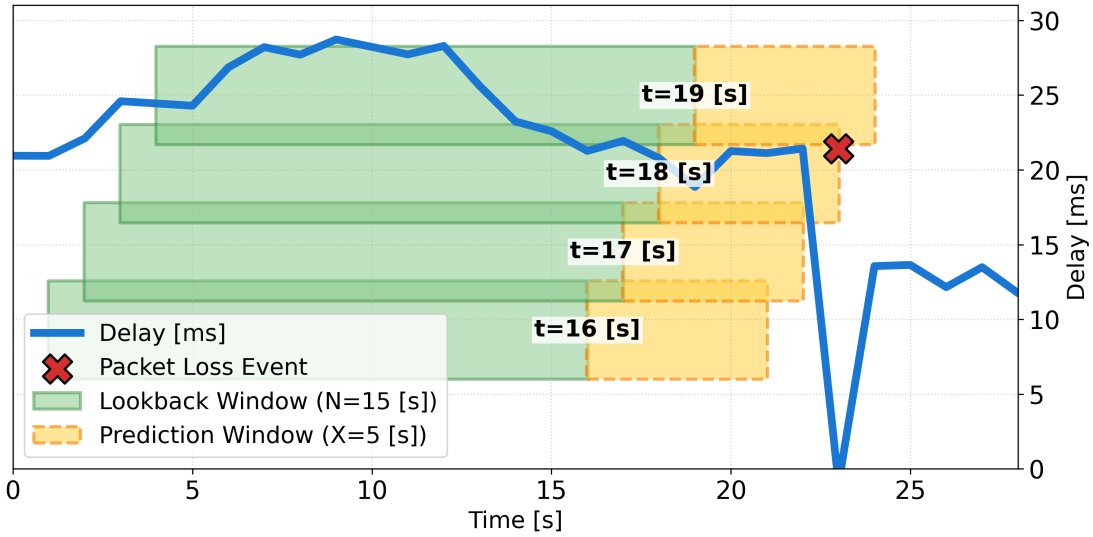


Evident **correlation** between packet loss and delay increase

Can we predict packet loss event before it happens looking at historical delay values?

Project #4-5 – Packet Loss Event Classification
Hints

- **Data preprocessing**
  - The rows with `delay_ms` set to `-1` (packet loss event) must be pre-processed
- **Procedure**
  - Use a sliding window approach: take the last ***N* seconds** of measurements (*Lookback Window*) as input to predict the presence of a packet loss event in the following ***X* seconds** (*Pred. Window*)
  - **Experiment with different *N* and *X* values**
  - For each Lookback Window statistical features can be used



Project #5 – Packet Loss Event Classification
Packet Loss Event Classification with XGBoost & Federated Learning

- (12 points) Perform classification on packet loss event using two different models
  - Neural Network
  - XGBoost
  - Evaluate feature importance
  - Experiment with different N and X values (see previous slide)
- (3 points) Advanced Task: Federated Learning
  - Assume a centralized entity (e.g., SD-WAN Controller) periodically collects local models (e.g., trained with data available at individual CPEs), aggregates them and redistribute the information to build a generalized model
  - Compare the performance between the general (i.e., federated) and the CPE-specific (local) models