A World Leading SFI Research Centre

Trinity College Dubli
Coláiste na Tríonóide, Baile Átha Clia
The University of Dublin

CONNECT
Centre for Future Networks

Compute and networking for edge cloud
Is AI both the problem and the solution?

Prof. Marco Ruffini
Dept. Computer Science and Statistics, Trinity College Dublin
CONNECT and IPIC research centres

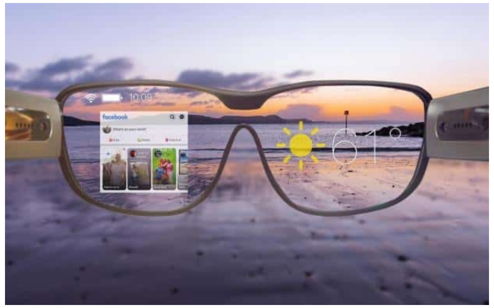
Trinity College Dublin
The University of Dublin
Ireland's European Structural and Investment Funds Programmes 2014-2020
Co-funded by the Irish Government and the European Union
European Union
European Regional Development Fund
Science Foundation Ireland For what's next
IPIC
BRINGING PHOTONICS TO LIFE

Connect — Networks of the Future

Trinity College Dublin

High performance VR today

- There is a large amount of computation, for which you need either external support… (cabled device)
- Or can do without PC and cable, sacrificing some performance
- <u>Or wait for this…</u>



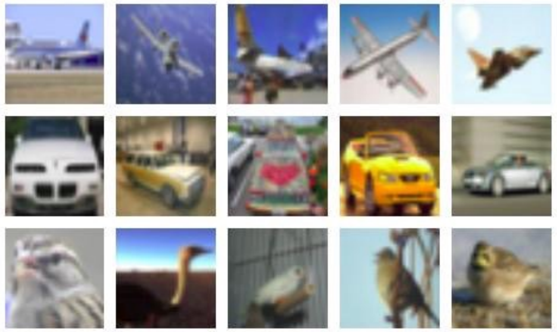
Object recognition

<u>…and offload your computation elsewhere</u>

Offloading to the edge

There is much research happening now on removing heavy computation altogether, i.e. offloading computation to the edge.

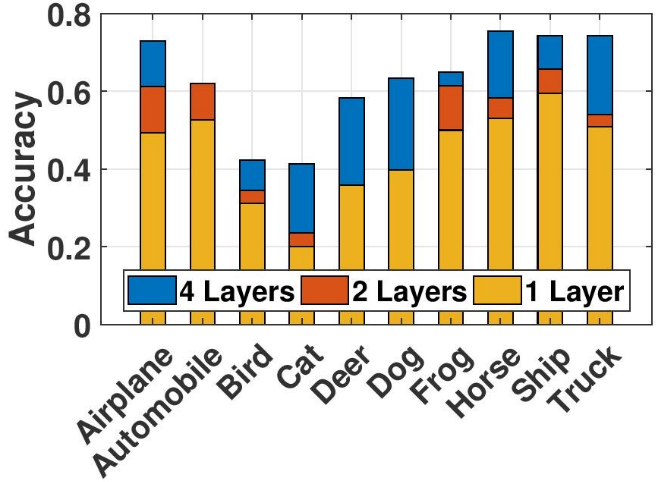
- Example of object recognition



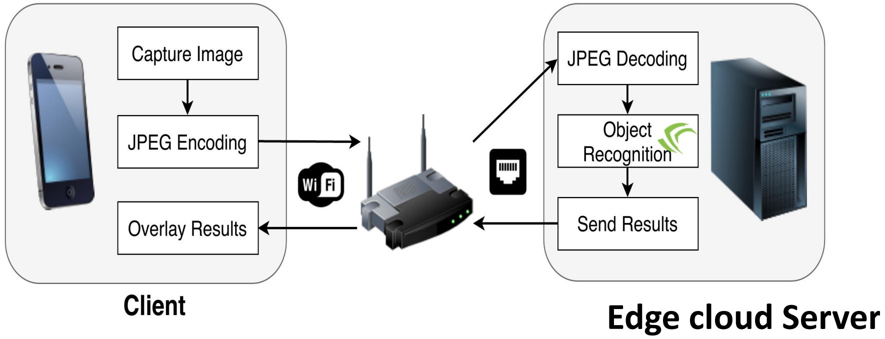
VS

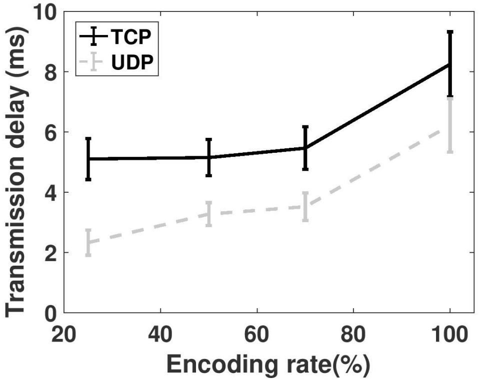
laptop

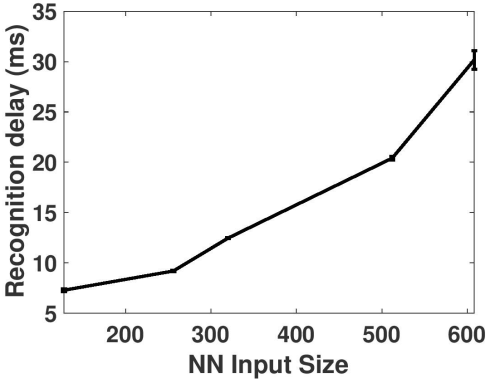
Average time: 2.537 s vs. 0.191

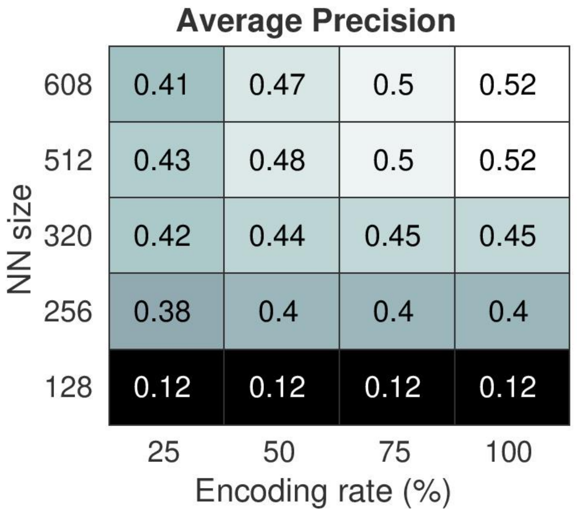


A. Galanopoulos, et al. Improving IoT Analytics through Selective Edge Execution, in proc. of IEEE ICC, 2020





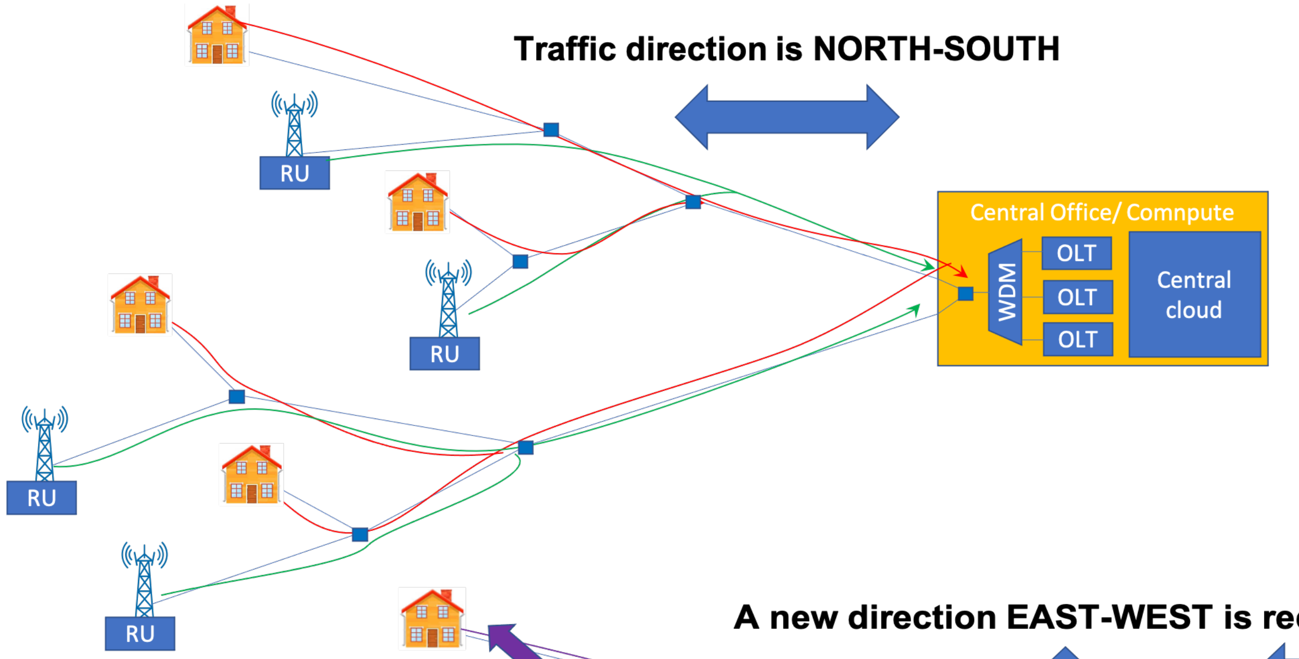




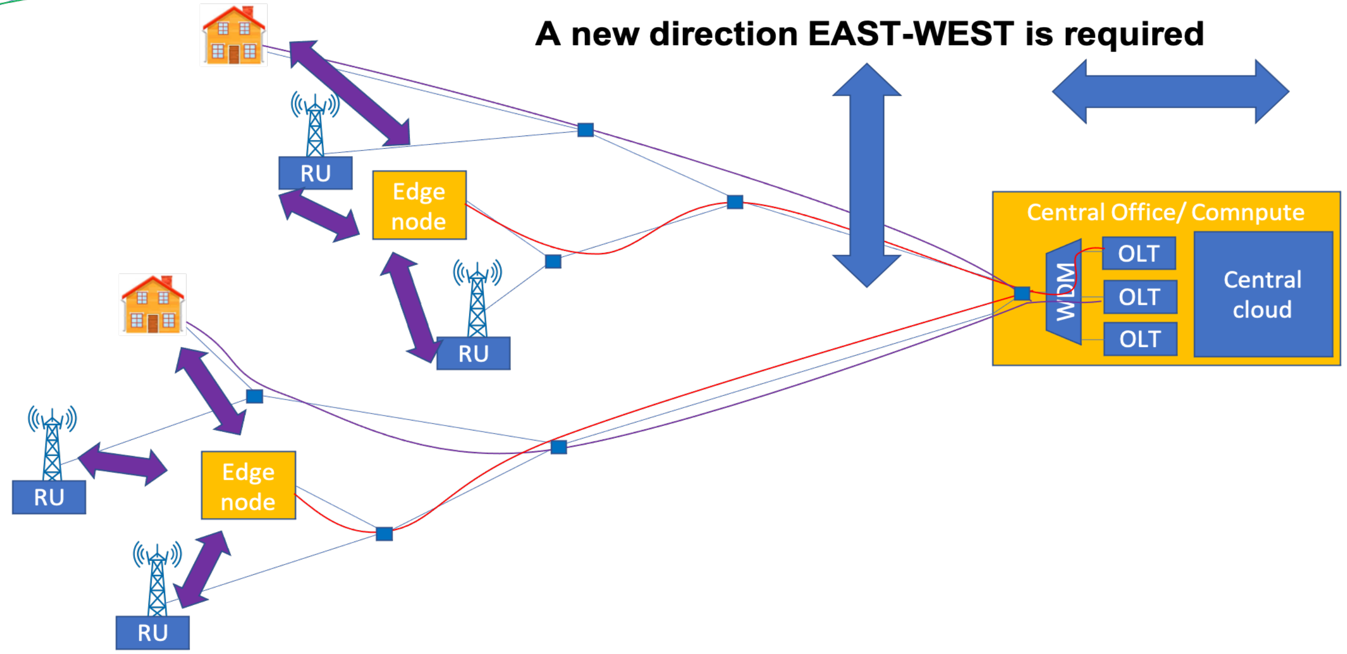
A. Galanopoulos, et al. Measurement-driven Analysis of an Edge-Assisted Object Recognition System, in proc. of IEEE ICC, 2020

The edge connectivity problem

- PON are widespread for high-performance, low-cost access connectivity.



- For lower latency there are MEC nodes… that's why they were invented
- But traffic to edge nodes requires handling of direct end points communications (EAST-WEST)
- <u>This is also crucial for mobile functional split</u>



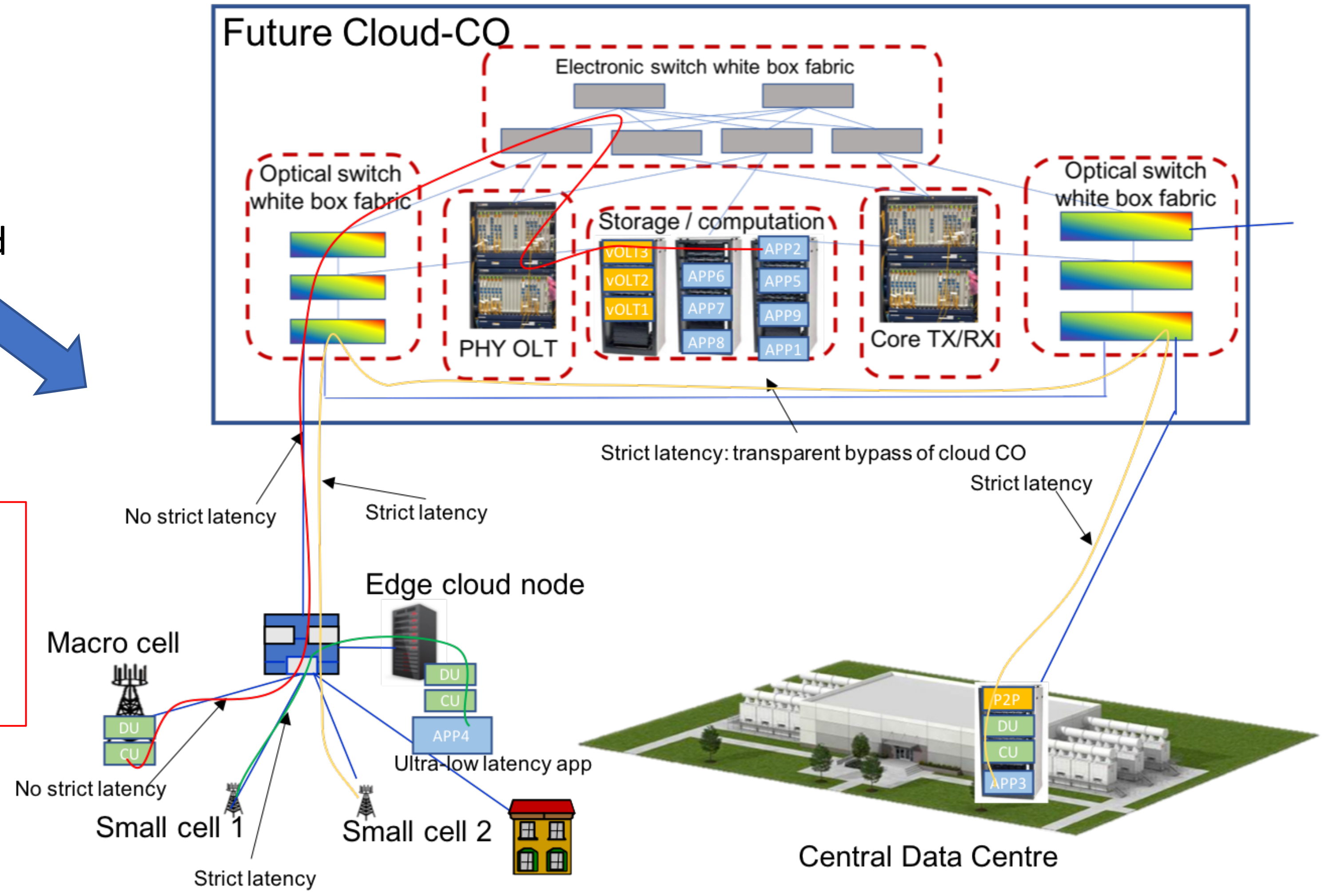
The statistical multiplexing problem

Demand varies in time and location:
→ take advantage of statistical multiplexing of compute capacity

→ This translates into any computing device should be reachable from anywhere

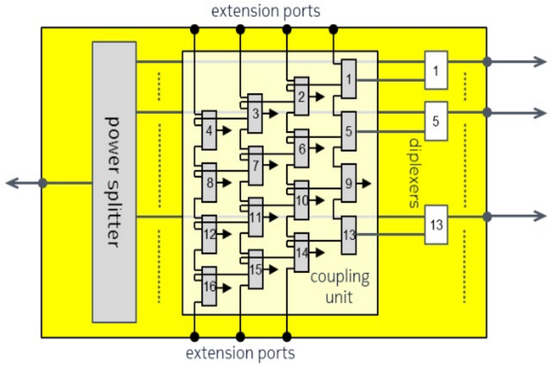
→ This brings up the network capacity and latency problem.

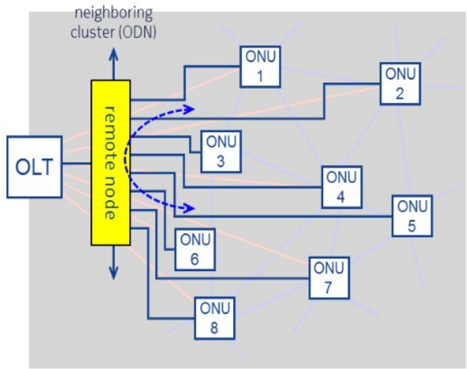
Dynamic optical connectivity can help on latency trade-off between distance (propagation time) and electronic bypass (routing and computation time).



Mesh-PON for connectivity

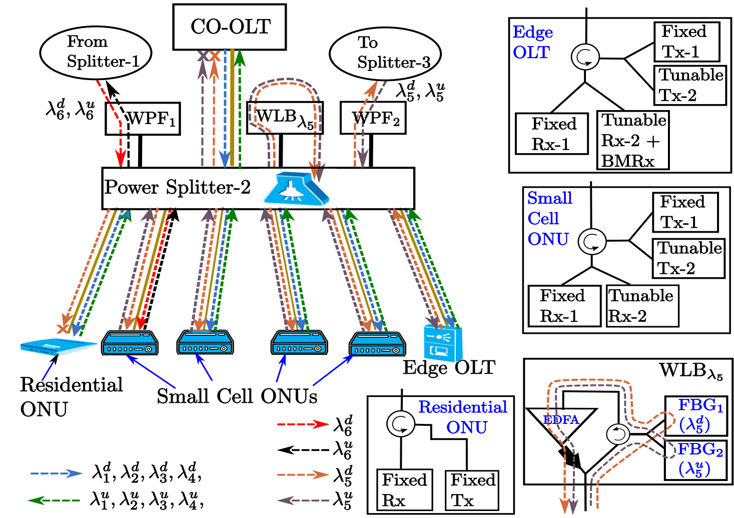
Th. Pfeiffer, "Converged heterogeneous optical metro-access networks," ECOC 2010, paper Tu.5.B.1





- S. Das, M. Ruffini. PON Virtualisation with EAST-WEST Communications for Low-Latency Converged Multi-Access Edge Computing (MEC). OSA Optical Fiber Communications Conference (OFC), March 2020

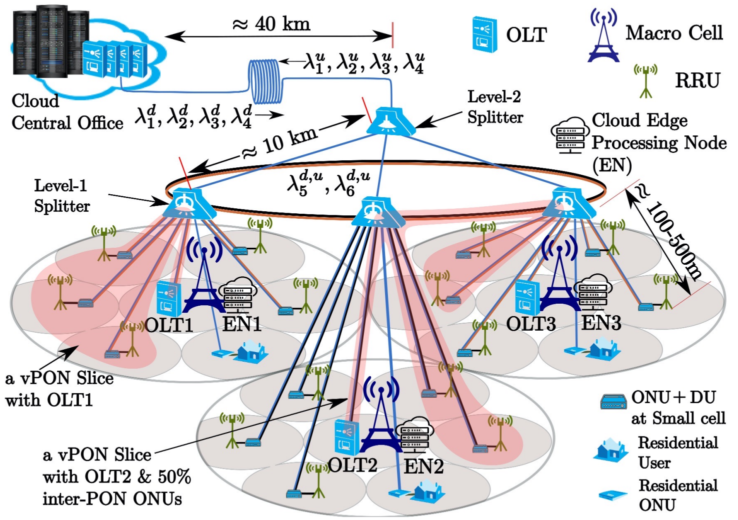
Dynamic solutions are more flexible and efficient but require power = this using reflections at splitter needs optical amplification.



Static solutions are less flexible (i.e., configuration of end points) and efficient (i.e., capacity usage) but can be totally passive.

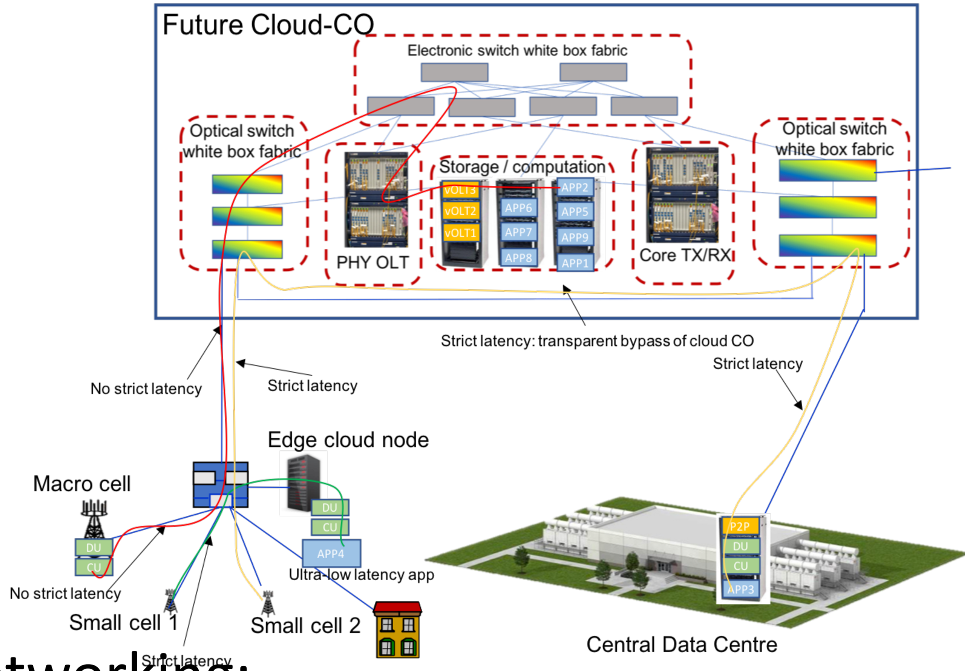
Here new technology at splitter nodes could be a game changer

- Whatever the splitter technology, virtualisation will play a defining role:
  - Dynamic organization/slicing of L2 scheduling across multiple wavelengths.
  - Organisation of wavelength routing in PON and across metro.
- Important constrained optimization problem…



AI for application but also resource allocation: it's both the problem and the solution

- AI computing is the problem: it sets the network and computing capacity and latency targets
- The solution is a complex dynamic optimization of <u>compute and network resource allocation</u>, <u>electronic and optical switching</u>, <u>transmission impairments.</u>
- Slicing and sharing adds inter-dependency between applications, requiring application/network performance monitoring.



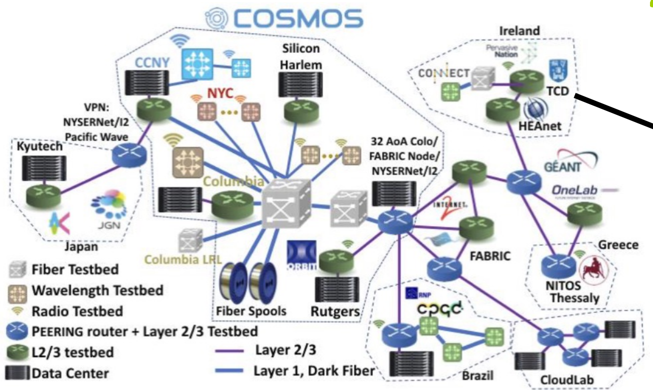
- AI is being applied to multiple aspects of computing and networking:
  - Demand and performance prediction
  - Estimating complex computing behavior, energy consumption, etc
  - Estimating optical transmission impairments

- Can application-level AI and network/compute optimization AI cooperate rather than adopting client/sever approach?

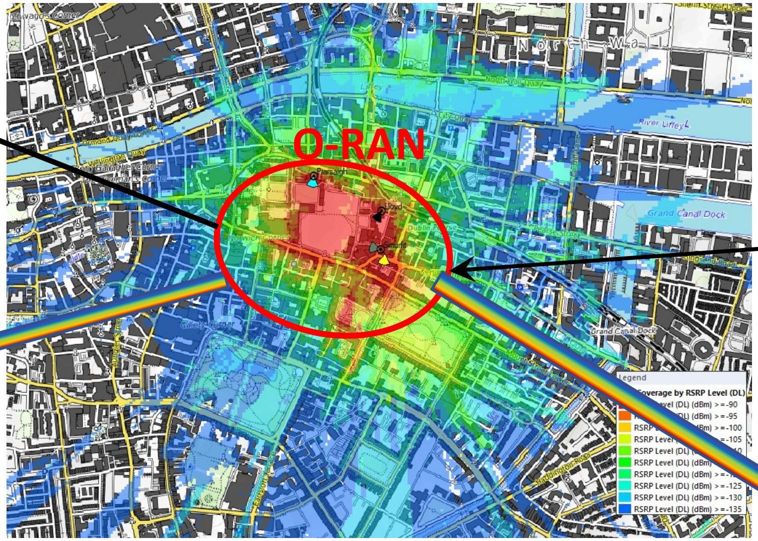
CONNECT Networks of the Future

TRINITY COLLEGE DUBLIN

Whatever you do: AI needs data and testing!

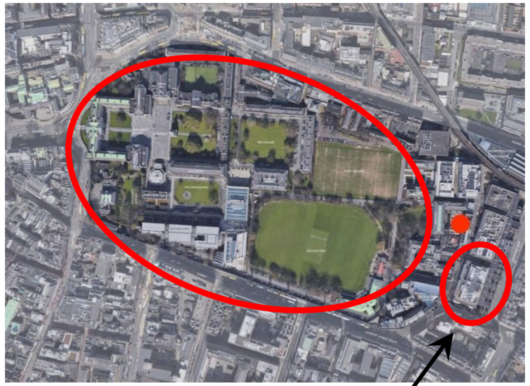


https://wiki.cosmos-lab.org/wiki



www.openireland.eu

Based in Trinity College campus



Radio and Optical Communications Laboratory

O-RAN

CONNECT research centre building

Reconfigurable optical fibre layer: create any topology across wireless, optical computing

Wireless based on **OpenRAN** (outdoor RU+ DU-CU) and **Opensource 5G** (using indoor USRP RF hardware)

Reconfigurable **Open Optical** system including:

- Full metro transmission: 2,000 km fibre, ROADMS, Coherent transponders, in-line EDFAs
- Virtualised optical access: virtual sliceable PON, mesh PON architecture
- Laboratory measurements: OSA, scope, power meters, etc.

Computing: **severs anywhere** in the topology, P4 capability, etc.