IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

Energy-Efficient Orchestration and Placement of 5G Network Slices

PhD Student: Cecília **DE SOUSA LUZ ALMEIDA**

Supervisors:

Geraldine **TEXIER** (IRISA/IMT Atlantique)
Nicolas **HUIN** (IRISA/IMT Atlantique)
Christelle **CAILLOUET** (Université Côte d’Azur)

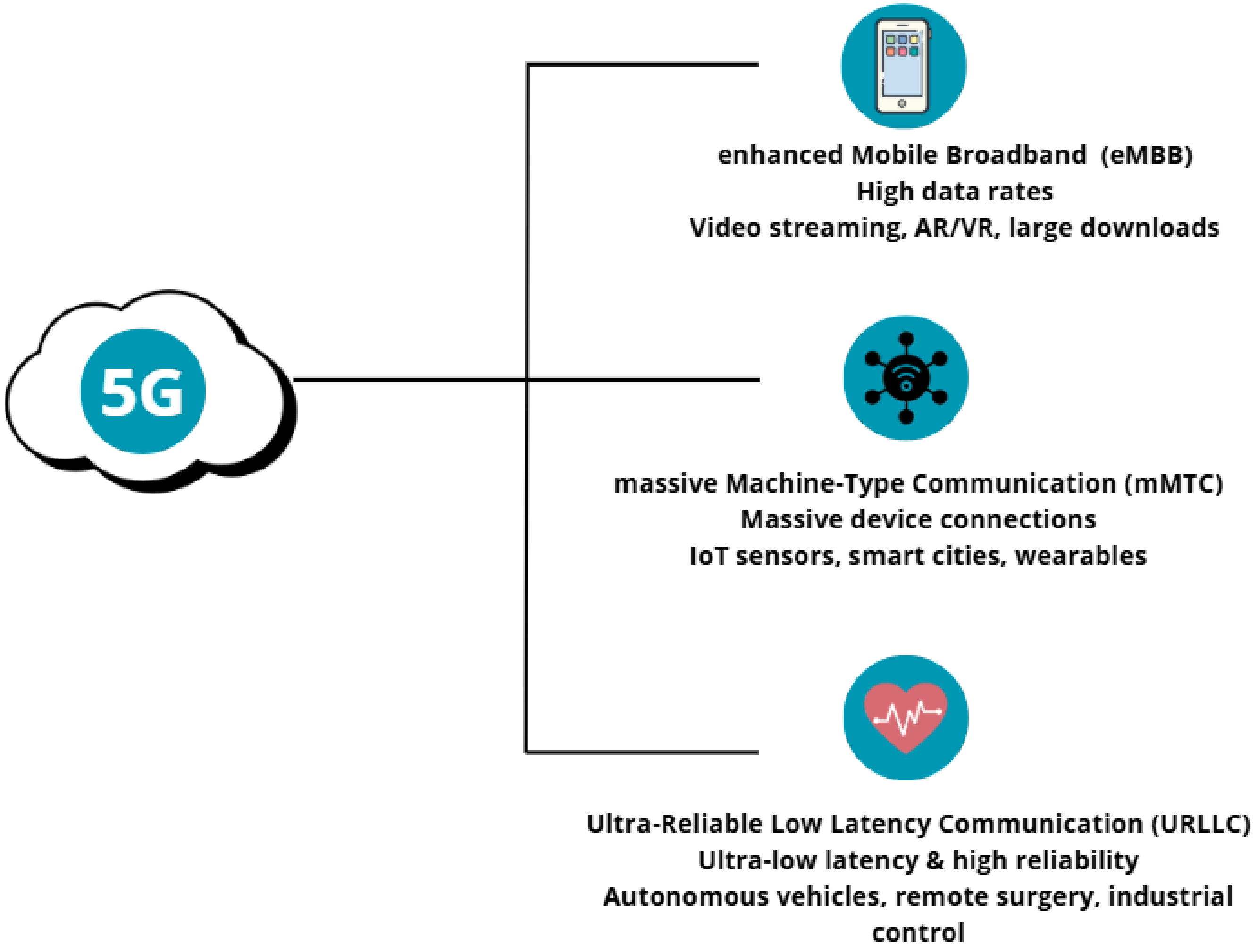
Financed by – PEPR Réseaux du Futur 2030

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

FRANCE 2030
PROGRAMME DE RECHERCHE
RÉSEAUX DU FUTUR

WHAT IS NETWORK SLICING?

- Multiple virtual networks on the same physical infrastructure;
- Each slice is tailored to different service requirements and strict SLAs;
- Although virtual, the slices consume and activate physical resource;



WHY ENERGY-EFFICIENT NETWORK SLICING MATTERS?

5G = many heterogeneous services
with different requirements

Slicing = more flexibility in how resources are used

Energy & carbon impact depend on how resources are used

Carbon & Energy Impact of 5G (France):
- **CO_2 footprint**: 16%–42% increase in CO_2 emission (2030)[1]
- **Power usage**: 16–40 TWh/year (2030)[1]

Main goal: *Optimize 5G slice placement to reduce energy use while ensuring service quality*

1.Source: Haut Conseil pour le Climat (2020). *5G and Climate – Impacts and Stakes of Mobile Telephony Evolution*.

STATE OF ART: ENERGY MODELING & METHODOLOGICAL DIMENSIONS

Dimension	Description / Categories
Modeling Scope	Core, RAN, Edge, or End-to-End architectures
Resource Abstraction Level	Physical (hardware-level) vs. Virtual/Logical (VNF or slice level)
Methodological Formulation	Empirical measurement vs. simulation-based modeling; Linear, Non-linear, or MILP optimization
Traffic & QoS Awareness	Consideration of bandwidth, latency, and service constraints

STATE OF ART: CATEGORIES OF ENERGY MODELS

Activation-based:
ON/OFF state;
minimize number of active elements.[1]

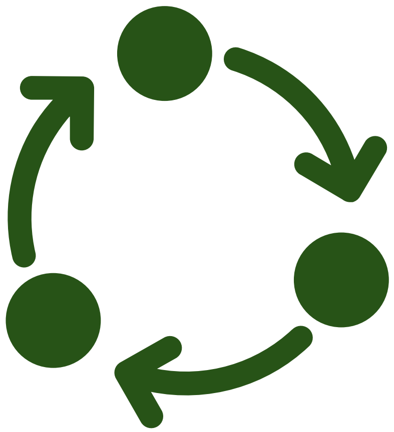
Load-dependent:
Power = Static + α·Load;
fine-grained estimation[3].

Efficiency-ratio:
Performance per unit energy (e.g., bits/Joule).[2]

1; Y. Magnouche, J. Leguay and F. Zeng, "Safe Routing in Energy-aware IP networks," *2023 19th International Conference on the Design of Reliable Communication Networks (DRCN)*, Vilanova i la Geltru, Spain, 2023, pp. 1-8, doi: 10.1109/DRCN57075.2023.10108152.
2. Y. Azimi, S. Yousefi, H. Kalbkhani, and T. Kunz, "*Energy-efficient deep reinforce-ment learning assisted resource allocation for 5g-ran slicing*," IEEE Transactions onVehicular Technology, vol. 71, pp. 856–871, 1 2022.
3. M. Elkael, A. Araldo, S. D'Oro, H. Castel-Taleb, M. A. Aba, and B. Jouaber, "*Joint placement, routing and dimensioning at the network edge for energy minimization*,"in Proceedings - IEEE Global Communications Conference, GLOBECOM. Instituteof Electrical and Electronics Engineers Inc., 2023, pp. 941–946.

STATE OF ART: ROUTING LITERATURE GAP

- Most energy-aware works focus on slice configuration and/or customization (e.g., placement, resource sizing, SLA satisfaction, low-consumption VNFs or slices)
- Routing is usually fixed or treated only as a constraint
- As a result, the dominant energy component of the network (links)is not explicitly optimized

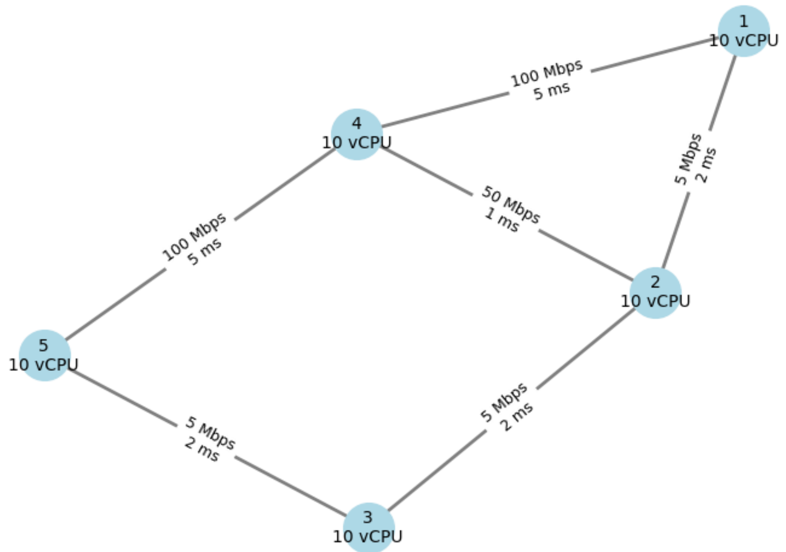


PLAN

- Problem Statement
  - Energy Model
  - Allocation & Routing Strategies
- MILP Formulation For Energy-aware Slice Allocation
- Preliminary Results
- Perspectives

PROBLEM STATEMENT

- Based on TAGHAVIAN (2024)[1] in VNF placement in 5G networks:
  - Routing-aware allocation
  - SLA constraints (bandwidth + latency)
- **Topology & Slice Model**
  - **Nodes:** homogeneous units with fixed CPU (host VNFs)
  - **Links:** undirected, with latency & bandwidth
  - **Slices:** chains of VNFs with CPU and throughput needs
  - **Virtual Links:** bandwidth + latency constraints



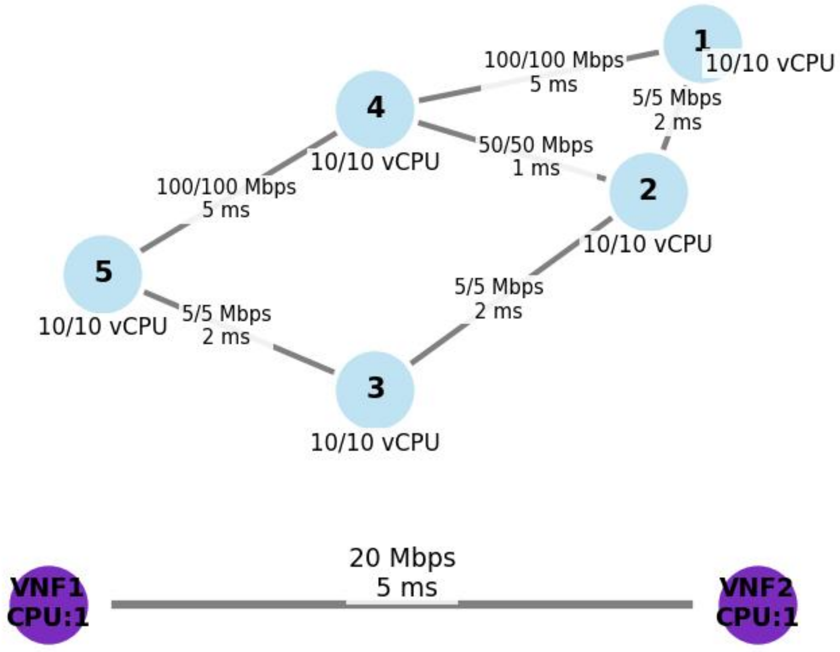
Example of Network Topology

Example of a Slice with 2 VNFs and 1 VL

1. **Masoud TAGHAVIAN**, *VNF Placement in 5G Networks using AI/ML,* PhD thesis, IMT Atlantique, 2024.

PROBLEM STATEMENT

- The problem consists in:
  - placing each VNF of the slice on a physical node
  - routing each virtual link between VNFs over the physical network



PROBLEM STATEMENT : ENERGY MODEL

- Energy Model Chosen: *Load-Dependent*
- **Node energy:**
  - Depends on **CPU usage**
  - Zero if the node hosts no VNF
  - C_n being the total capacity of the node

$$E_n^{\text{node}} = \begin{cases} 0, & \text{if no VNFs are placed,} \\ E_{\text{static}}^{\text{node}} + \alpha_n \cdot \text{CPU}_n^{\text{used}}, & \text{otherwise.} \end{cases} \qquad \text{with } \alpha_n = \frac{1}{C_n}$$

- **Link Energy:**
  - Depends on **bandwidth usage**
  - Zero when no traffic flows
  - BW_n being the total bandwidth of the link

$$E_l^{\text{link}} = \begin{cases} 0, & \text{if no links are placed,} \\ E_{\text{static}}^{\text{link}} + \beta_n \cdot \text{BW}_n^{\text{used}}, & \text{otherwise.} \end{cases} \qquad \text{with } \beta_n = \frac{1}{BW_n}$$

- **Total Energy:**

$$E_{\text{total}} = \sum_n E_{node} + \sum_\ell E_{link}$$

PROBLEM STATEMENT : ALLOCATION & ROUTING STRATEGIES

- Based on the algorithm proposed by TAGHAVIAN (2024), this work implements joint VNF placement and routing using A* and two other variants

A*:

Latency-based routing

- Classic pathfinding with heuristic with latency being the cost
- Finds feasible VNF-to-VNF routes
- Ensures chain connectivity

PROBLEM STATEMENT : ALLOCATION & ROUTING STRATEGIES

- Based on the algorithm proposed by TAGHAVIAN (2024), this work implements joint VNF placement and routing using A* and two other variants

A*:

Latency-based routing

- Classic pathfinding with heuristic with latency being the cost
- Finds feasible VNF-to-VNF routes
- Ensures chain connectivity

A*-Bandwidth (ABO):

Bandwidth-aware routing

- Prioritizes paths with higher residual bandwidth
- Reduces congestion and routing failures
- Better satisfies bandwidth SLA

PROBLEM STATEMENT : ALLOCATION & ROUTING STRATEGIES

- Based on the algorithm proposed by TAGHAVIAN (2024), this work implements joint VNF placement and routing using A* and two other variants

A*:

Latency-based routing

- Classic pathfinding with heuristic with latency being the cost
- Finds feasible VNF-to-VNF routes
- Ensures chain connectivity

A*-Bandwidth (ABO):

Bandwidth-aware routing

- Prioritizes paths with higher residual bandwidth
- Reduces congestion and routing failures
- Better satisfies bandwidth SLA

Fairness A*-Bandwidth (FABO):

Load-aware routing

- Adds load balancing to avoid overusing nodes/links
- Penalizes overused resources → spreads traffic more evenly
- Improves long-term network stability and fairness

PROBLEM STATEMENT : ALLOCATION & ROUTING STRATEGIES

- We also propose an **energy-aware extension of A*** that performs joint VNF placement and routing by minimizing incremental infrastructure energy.

Energy-Aware A*:

Energy-driven search

- Replaces latency-based cost with incremental energy cost:
  - Node activation energy
    (baseline cost + proportional CPU usage)
  - Link activation energy
    (baseline cost + proportional bandwidth usage)
- Considers node activation and link activation
- Accounts for CPU and bandwidth utilization
- Preserves feasibility constraints (CPU, BW, latency)

MILP FORMULATION FOR ENERGY-AWARE SLICE ALLOCATION

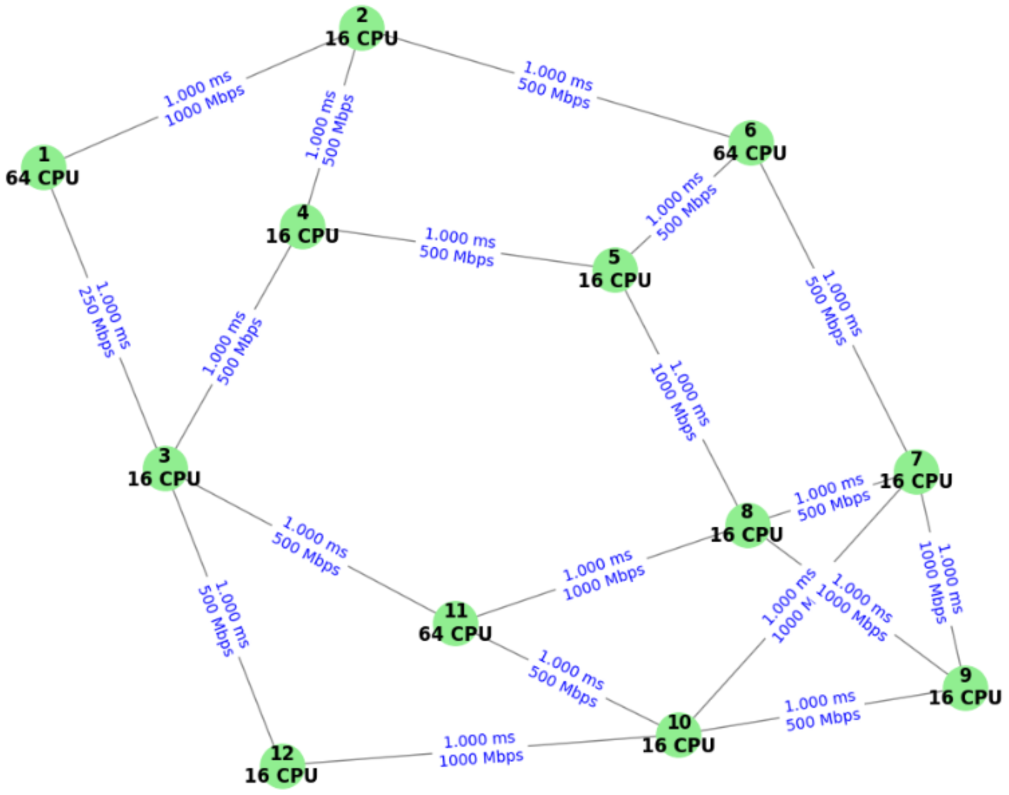
- **Goal:** Find the optimal **joint placement of VNFs** and **routing of virtual links**, minimizing **energy consumption** under **CPU, bandwidth, and latency constraints**.
- **Decision Variables**
  - **Placement**
    - $\mathrm{x}_{i,n} = 1$ if VNF i is placed on node n
  - **Routing**
    - $y_{i,j,u,v} = 1$ if virtual link (i,j) uses physical link (u,v)
  - **Node activation**
    - $a_n = 1$ if node n hosts at least one VNF
  - **Link activation**
    - $b_{u,v} = 1$ if physical link (u,v) carries traffic
- **Objective Function**
    - Phase 1: Maximize number of slices accepted:

$$Z^{\star} = \max \sum_{s \in \mathcal{S}} z_s$$

    - Phase 2: Minimize total energy:

$$[\min \sum_{n} a_n \left(1 + \frac{\mathrm{CPU}_n^{used}}{C_n}\right) + \sum_{(u,v)} b_{uv} \left(1 + \frac{\mathrm{BW}_{uv}^{used}}{B_{uv}}\right)]$$

PRELIMINARY RESULTS

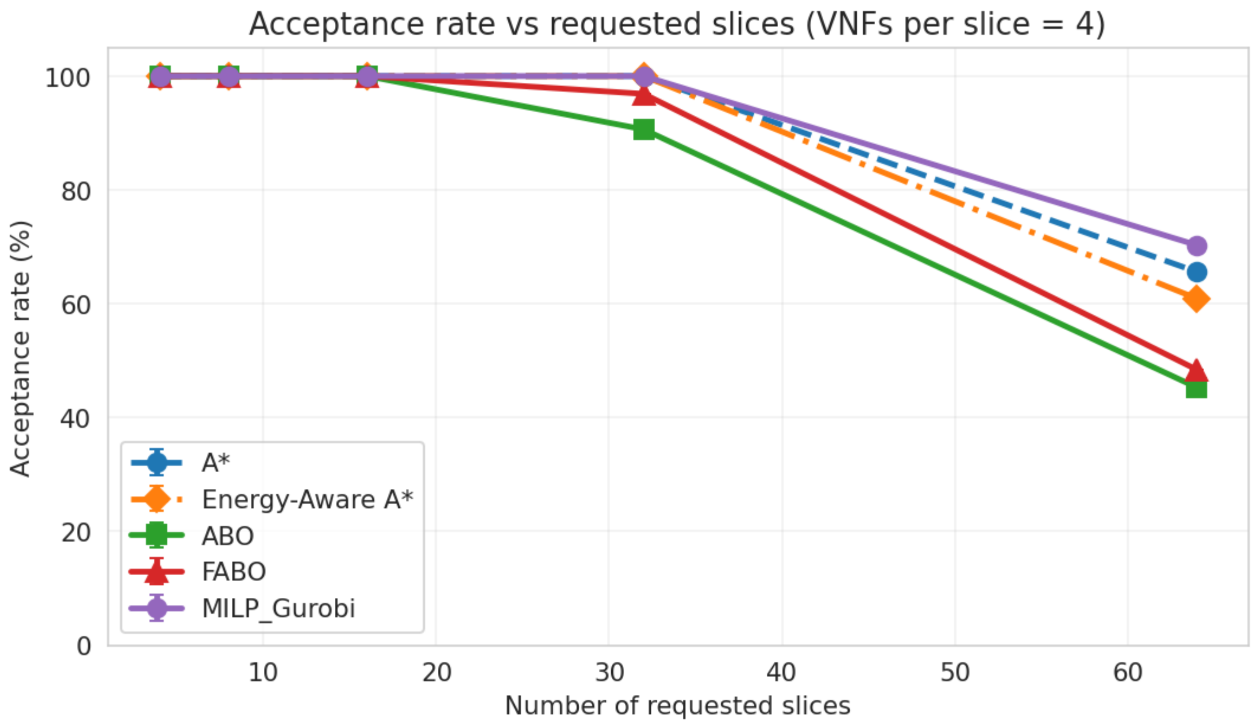


12-node topology inspired by a Finnish network

- 25 scenarios (5 random seeds):
  - 4 to 64 slices
  - VNF chains of 2–6 functions
- Evaluated metrics:
  - Acceptance rate
  - Total energy consumption vs load

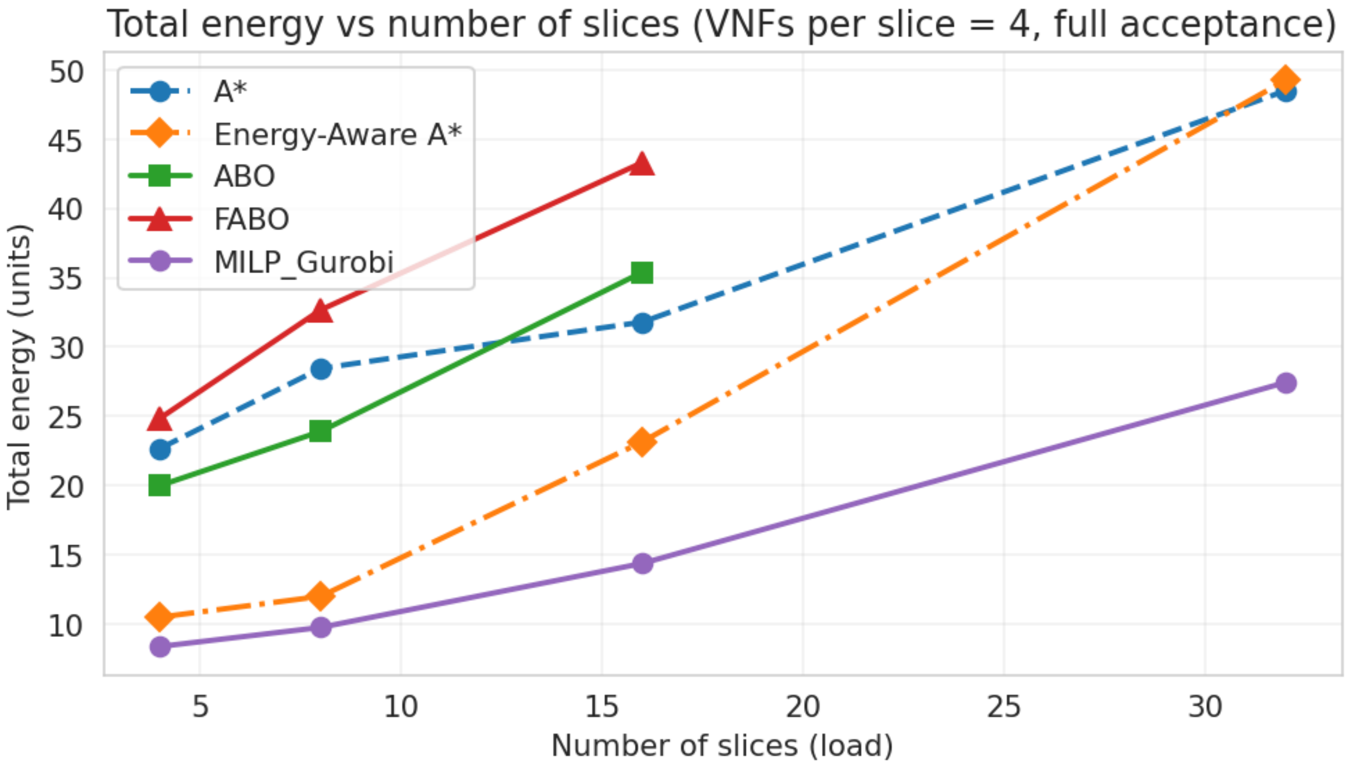
Extracted from: Lira, Rodrigo & Martins Filho, Joaquim & Bastos-Filho, Carmelo. (2018). *An adaptive–alternative restoration algorithm for optical networks*. Photonic Network Communications. 35. 10.1007/s11107-017-0729-5.

PRELIMINARY RESULTS: ACCEPTANCE RATE VS LOAD



- **Low load (≤16 slices):** all methods achieve 100% acceptance.
- **Congestion regime (32+ slices):** acceptance decreases for all methods. **MILP achieves the highest feasibility under high load**, serving as an upper bound.
- **A*** and **Energy-Aware A*** remain competitive, with similar scalability behavior.
- **ABO and FABO degrade earlier**, indicating higher sensitivity to bandwidth saturation.

PRELIMINARY RESULTS: TOTAL ENERGY CONSUMPTION VS LOAD



- **MILP** provides the best energy efficiency under full feasibility.
- **Energy-Aware A*** reduces unnecessary infrastructure activation.
- Path diversification strategies increase robustness but raise energy cost.
- Energy-performance trade-offs become more visible as load grows.

FUTURE PERSPECTIVES

- **Systematic scalability analysis**
  Evaluate heuristic sensitivity to load, congestion, resource imbalance, and SLA aspects that remain underexplored.
- **Routing–energy coupling analysis**
  Quantify how routing decisions impact global energy consumption across topologies
- **Robustness under heterogeneous demand**
  Assess stability under diverse traffic patterns and SLA constraints
- **MILP and heuristica refinement and benchmarking**
  Strengthen formulation consistency and use the MILP as a reference baseline to improve heuristic competitiveness at high load.