

AI-Driven Energy Efficiency for 6G Networks (2/2)

Ayat Zaki-Hindi (Luxemburg Institute of Science and Technology)

6G-EWOC, 6G-TWIN, DESIRE6G, EXIGENCE

EUCNC 2025 Workshop on “Technology Enablers for Sustainable 6G Design”

June 3, 2025

Overview

- Personal Health & Protection from Harm
- Environmental Sustainability
- Federated cross-domain radio and computing resource optimisation
- Optimisation of 6G radio access technologies
- Programmable and AI-Enhanced SDN for Energy-Efficient 6G
- Digital Twins as Enablers of Sustainable 6G Operations
- AI-Native Task Redirection for Sustainable Execution
- Energy-Efficient FL at deep edge
- Carbon-aware AI Services



Personal Health & Protection from Harm

Context:

- **Road safety** is a primary concern as accidents cut short the lives of approximately 1.2 million people every year, and responsible for countless non-fatal injuries, many of them incurring disability.

6G-EWOC approach

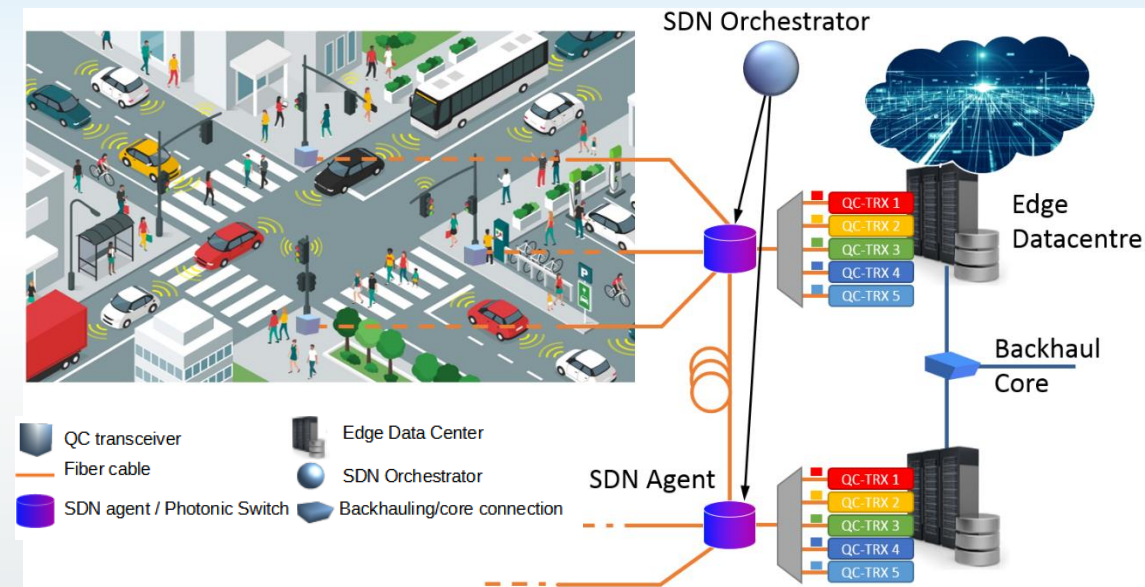
- **Connected and automated driving**, enabled through **instantaneous access to information** for sharpening the situational awareness, can mitigate this toll on our society while enhancing the efficiency for transporting humans and goods.
- **Large volume of information to be shared** and made available to all traffic participants.
- Inclusion of **precise sensors, connectivity at low latency**, and a powerful compute infrastructure to **fuse**, in real time, the **vast amounts of data generated** along the roadside scenery.



Environmental Sustainability

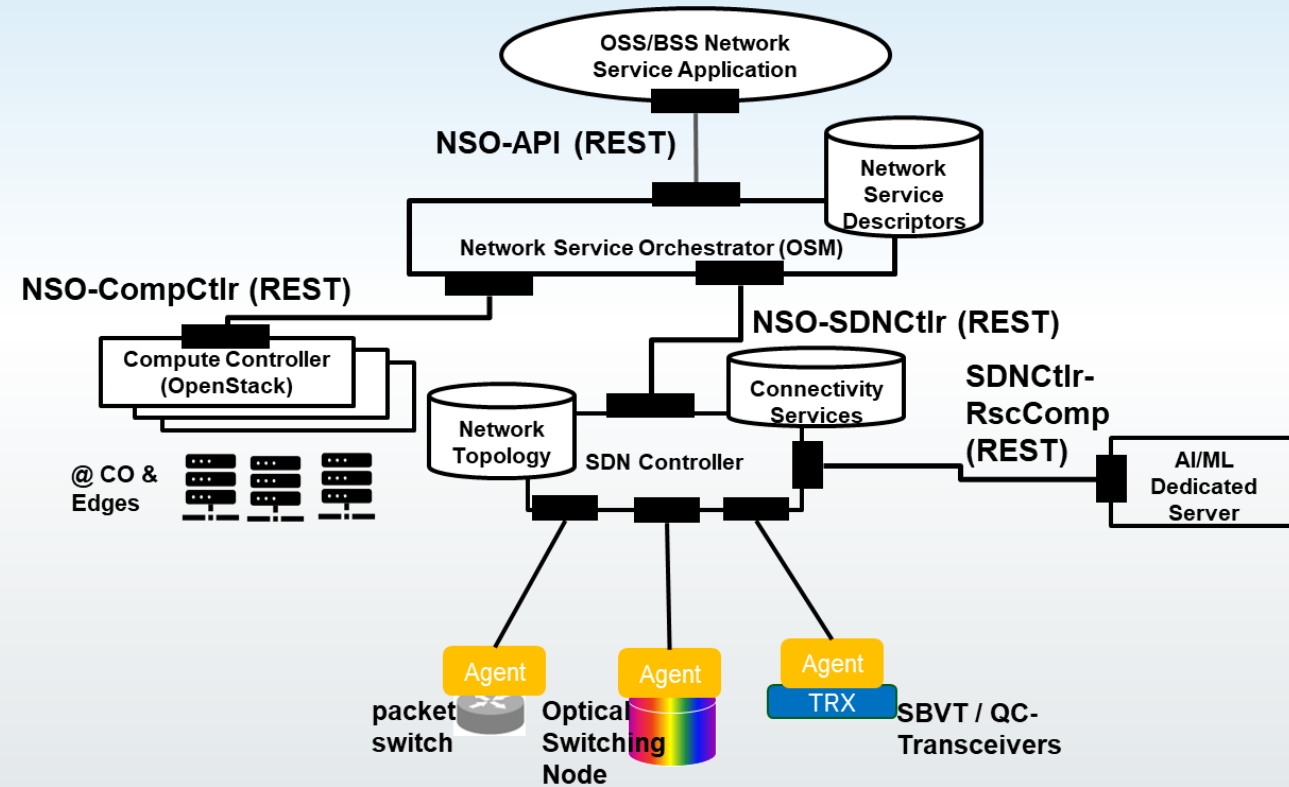
Context:

- “... deployment of 6G, environmental sustainability involves promoting efficient resource utilization, reduction in energy emission, reduction of carbon footprint, and mitigating other negative environmental impacts and rebound effects” [1]
- **Large volume of information to be shared** and made available to all traffic participants... and reducing energy consumption.
- 2 KPIs to address environmental sustainability:
 - Provisioning of traffic flows achieving 50% reduction of the energy consumption
 - AI-assisted energy-efficiency algorithm(s) and/or heuristics for multi-layer (packet/optical) networks.



6G-EWOC approach

- **AI-enhanced SDN** for dynamic traffic optimization, achieving significant energy savings and rapid provisioning across multi-layer networks
- **Abstraction Models** between hierarchical control entities, supported autonomous operations (creation/updating/terminating), monitored information, etc.
- **Data Models for programming underlying optical devices** such as Node and Transceivers
- **Devising AI/ML models** (e.g., exploiting DRL) to support *multi-objective path and resource selection*: efficient use of network resources and energy-efficiency – Delegated function of the SDN Control.



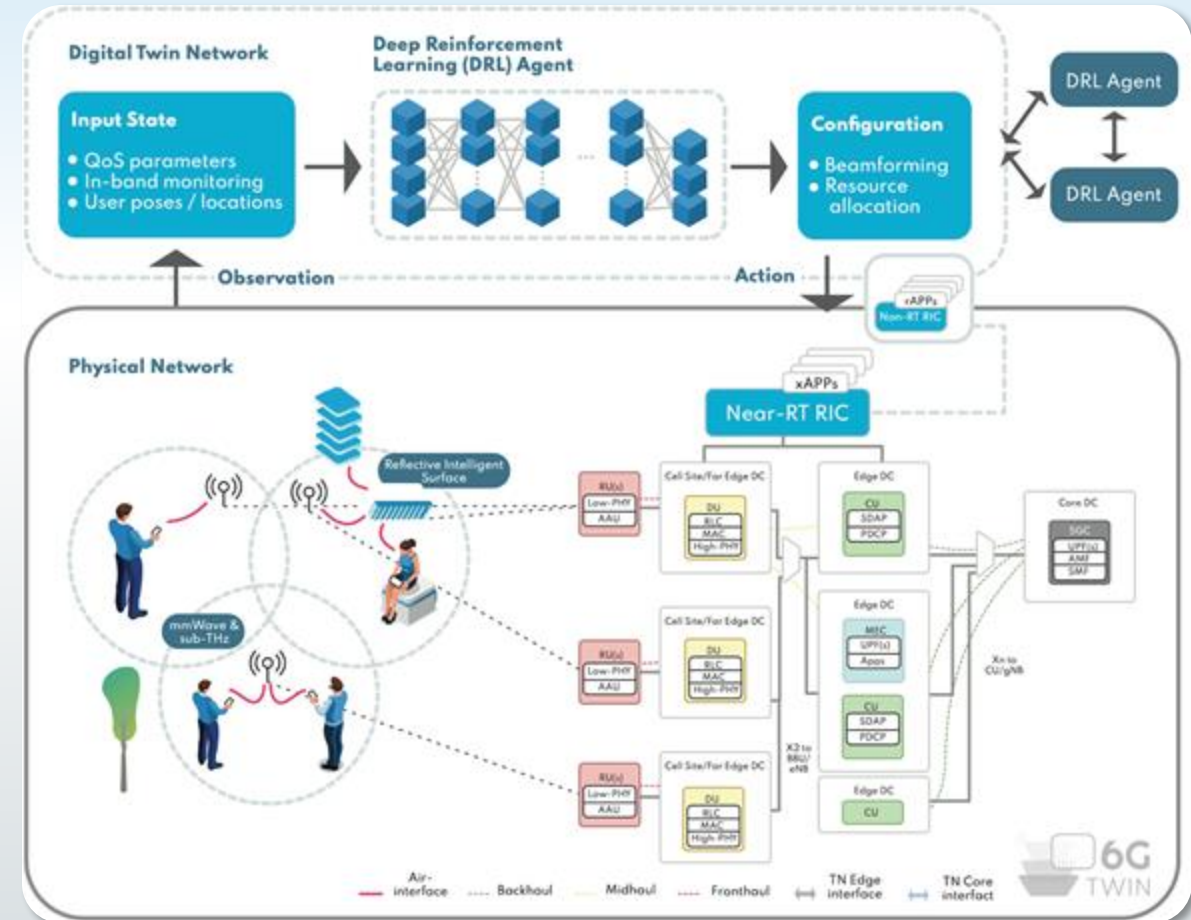
Federated Cross-domain Radio And Computing Resource Optimisation

Context

- **Advanced sleep mode** for **RAN** is key technique to reduce the energy consumption
- From a **MEC** perspective, activating or deactivating a base station translates into **variations in the aggregated traffic** that must be processed
- RAN and Edge operate under separate administrative domains, requiring a **federation** mechanism to achieve end-to-end energy efficiency

6G-TWIN approach

- Building **NDT RAN** and **Edge** components from the collected data from the infrastructure
- The system simulates different “**what-if**” scenarios, testing multiple energy efficiency strategies before deployment on the real-world network
- The NDT **recommends** the most suitable strategy for each tested scenario



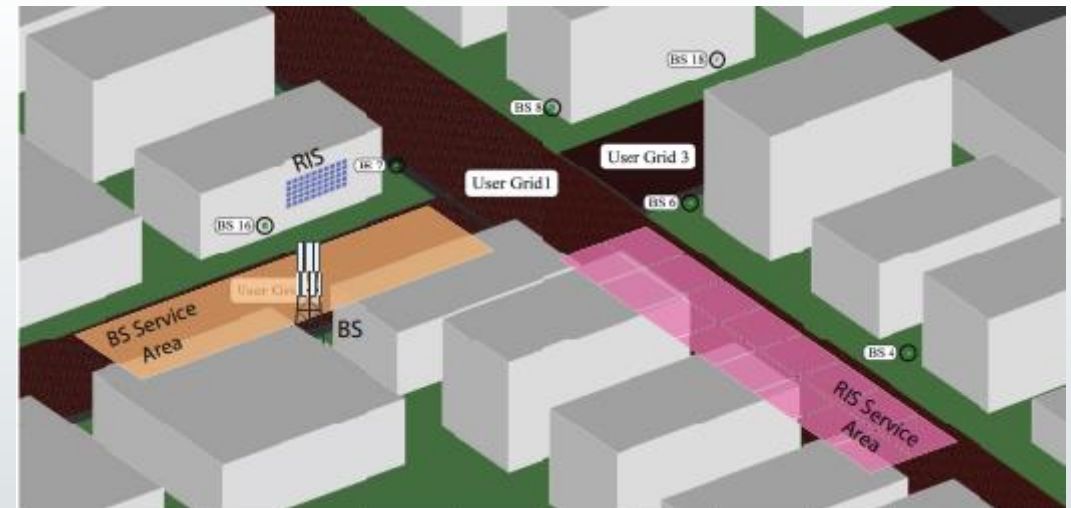
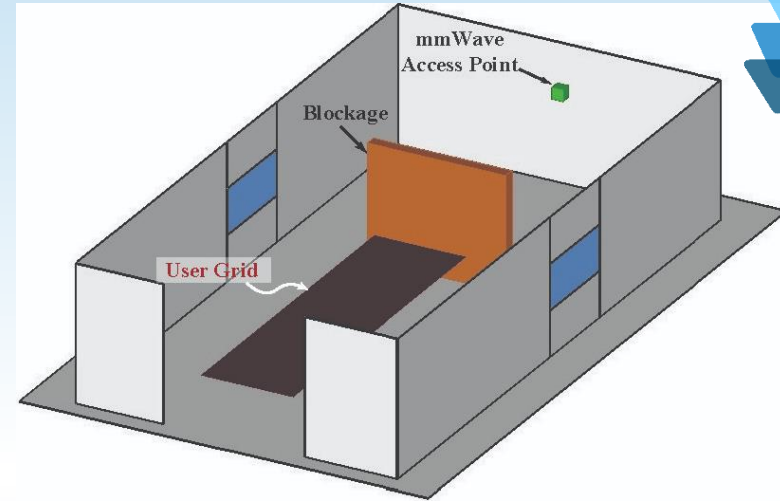
Optimisation of 6G Radio Access Technologies

Context

- **Higher frequencies**, such as mmWave and sub-THz, combined with **Reconfigurable Intelligent Surfaces (RIS)** are considered as an enabler for green communications
- Beam sweeping used for **beam management** can lead to big overhead in real networks

6G-TWIN approach

- Developing data-driven algorithms trained using **synthetic data** generated by the **environment-aware NDT**
- Expected enhanced **energy efficiency** by increasing the **accuracy of beam selection**, since with optimal beams less transmission power can be allocated to fulfil the application requirements



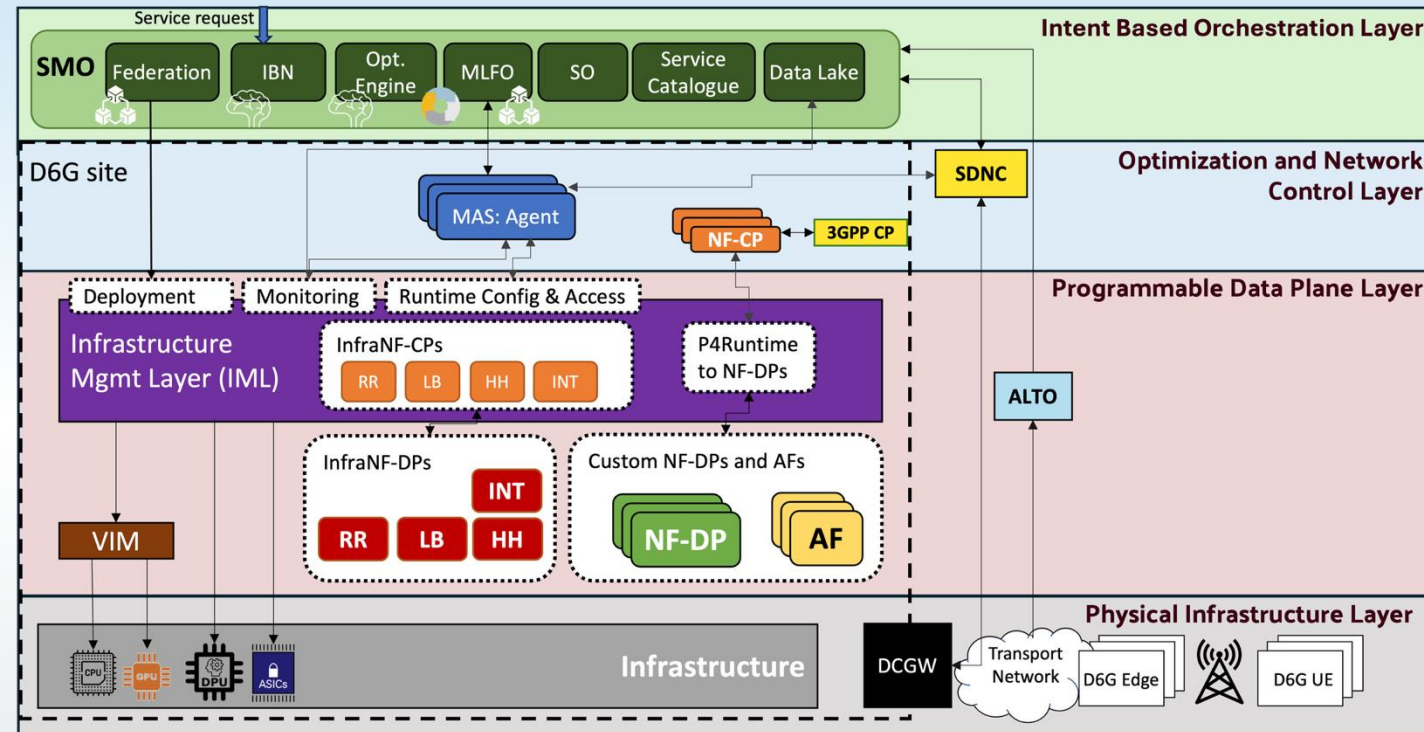
Programmable and AI-Enhanced SDN for Energy-Efficient 6G

Context

- AR UC – dynamic orchestration supports low-latency, energy-efficient streaming

DESIRE6G approach

- DESIRE6G architecture is enhanced with a multi-agent system
- **Multi-Agent Systems (MAS)** enable real-time, AI-based decision-making across the network-cloud continuum
- **In-band telemetry and dynamic traffic steering** provide adaptive traffic management, reducing overprovisioning
- **Cloud-native, Kubernetes-based architecture** supports modular, reusable deployments
- Soft prototypes show **promising energy savings through AI-driven self-optimization**



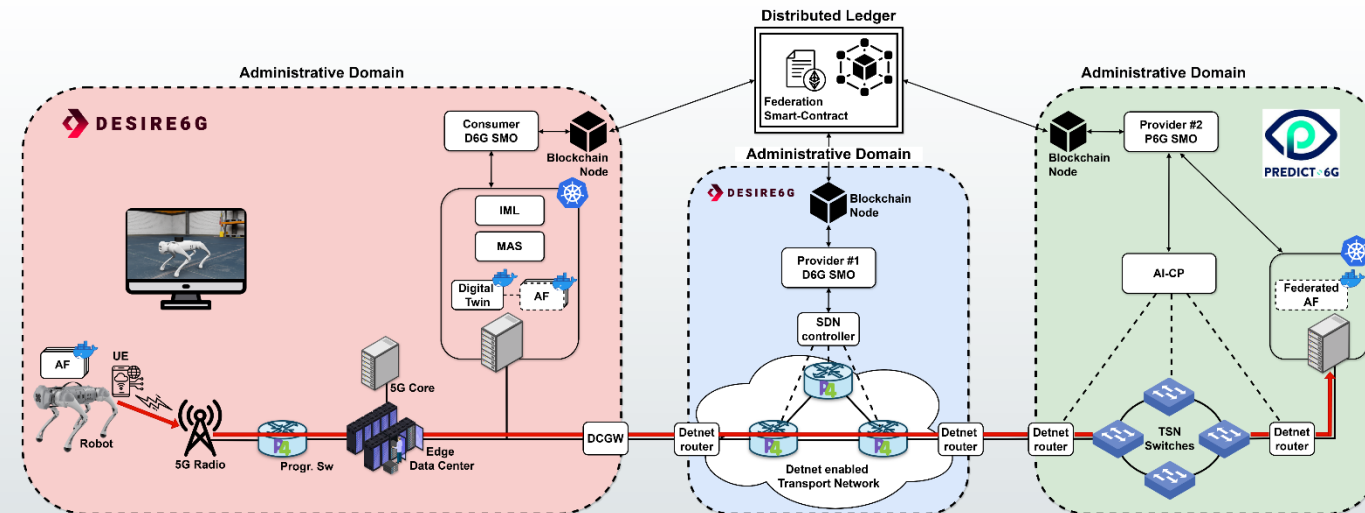
Digital Twins as Enablers of Sustainable 6G Operations

Context

- DT UC – robotic dog with ultra-low latency and predictive analytics for task control

DESIRE6G approach

- Real-time data and AI-driven **feedback loops** power predictive maintenance and resource optimization
- Digital replicas (e.g., robotic dog) maintain performance via sensor-driven, low-latency **control loops**
- Enables granular service lifecycle management with **energy-aware compute and communication**
- Programmable infrastructure supports **multi-level optimization** to meet KPIs efficiently



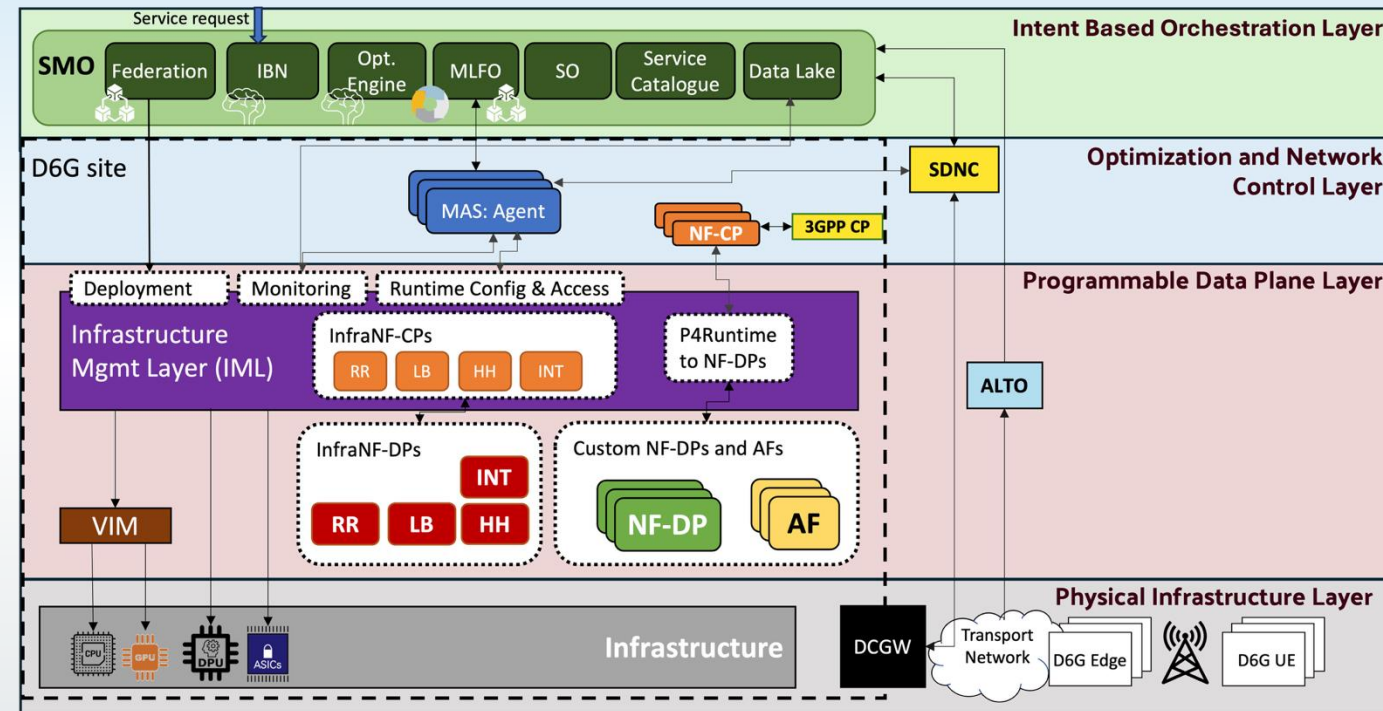
AI-Native Task Redirection for Sustainable Execution

Context

- AR/DT UCs – offloading minimizes endpoint energy use while ensuring performance

DESIRE6G approach

- **SOL stack** (scalable optimised learning) executes AI models efficiently across heterogeneous resources, minimizing energy usage
- **Dynamic task scheduling** redirects workloads to eco-powered or underutilized nodes
- **AI-native orchestration** supports real-time, energy-aware decisions
- Integrates with **programmable infrastructure** to maximize green resource usage



Energy-Efficient FL at deep edge

Context

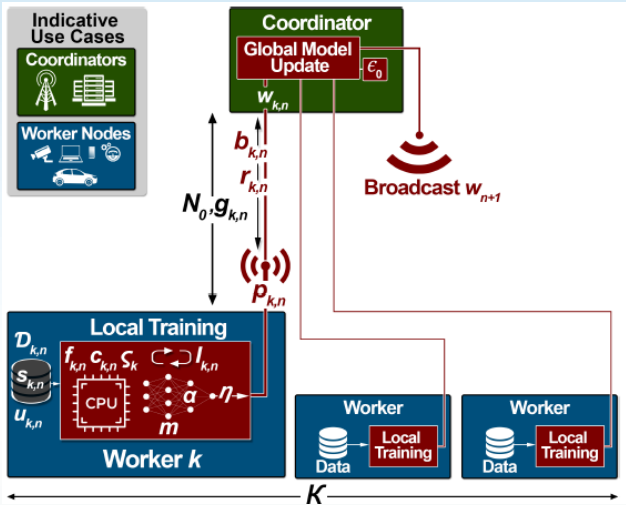
- To enable energy-efficient 6G networks powered by AI, the energy cost of running AI itself must also be minimized.
- **A federated learning system:**
 - A coordinator (residing at BS)
 - K workers of heterogenous comm. and comp. powers
 - **Goal:** find optimal set of trans. (p_k) and comp. capacity (f_k) at devices to minimize learning energy

EXIGENCE Approach [1]

- A **Soft Actor-Critic (SAC)** agent is trained to select p_k and f_k in real-time
- A **penalty-aware reward mechanism** to fulfil energy and time constraints and prevents inefficient strategies
- **Lightweight simulation** for offline training, enabling rapid deployment with minimal resource overhead

Impact

- Up to 94% reduction in total energy consumption compared to SOTA
- Scalable to dynamic wireless environments and varying AI task complexities



An FL system with one coordinator and K workers

Total Avg. (±STD)	5 Workers					10 Workers				
Schedulers	SAC	GAS	BES	RSS	GSS	SAC	GAS	BES	RSS	GSS
Total	35.1 (±12.6)	53.1 (±20.8)	603.5 (±172.3)	237.7 (±74)	148 (±54.5)	86.6 (±21.7)	118.7 (±26.1)	1343.7 (±250)	542.5 (±103.4)	465 (±95.8)
Computation Energy (J)	19.5 (±8.3)	29.1 (±15.1)	324.7 (±92.4)	107.5 (±34.1)	61.5 (±29.5)	59 (±16.6)	71.8 (±17.5)	724.1 (±143.6)	245.2 (±53.6)	285.4 (±67.8)
Transmission Energy (J)	15.6 (±7.8)	24 (±8.4)	278.8 (±100.4)	130.2 (±49.5)	86.5 (±33.9)	27.6 (±9.7)	46.9 (±11.4)	619.6 (±142)	297.3 (±68.1)	179.6 (±62.2)
Training Time per Global Iteration (s)	8.5 (±2.1)	7.6 (±1.6)	3.7 (±0.7)	8.3 (±1.1)	8.2 (±2.4)	8.4 (±1.7)	7 (±1.1)	4.3 (±0.6)	10.3 (±0.9)	7.3 (±2.2)
Global Iterations	15.8 (±3.7)	15.4 (±3.8)	15.5 (±4)	15.3 (±3.9)	15.6 (±4)	17.9 (±2.7)	17.5 (±2.6)	17.3 (±2.6)	17.6 (±2.4)	18 (±3.1)

Up to **20 times!** reduction in energy consumption of the system without any reduction of accuracy

Carbon-aware AI Services

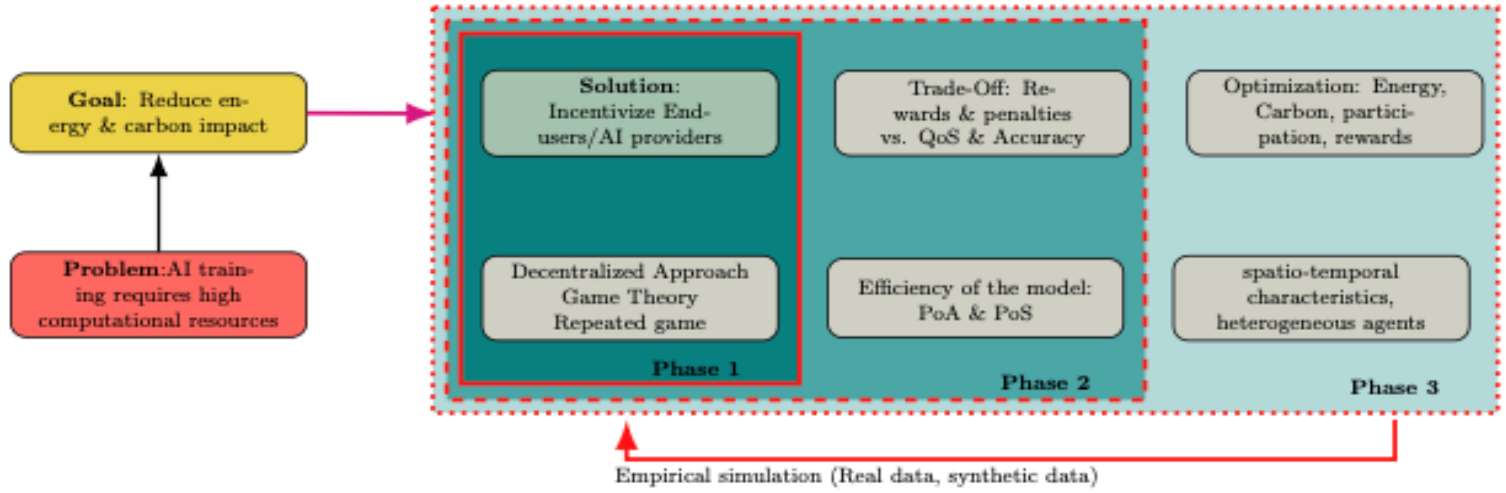


Context

- AI services require high computational resources for both training and inference
- Optimization of mechanisms (in terms of energy efficiency) can significantly reduce energy and carbon impact

EXIGENCE Approach

- Incentivize all related **Stakeholders** considering both business relationships and energy dependencies
- Incentive **Strategies for End-users** (individuals and enterprises / institutions):
 - Time-shifted execution, peer trading, behavioural rewards, discounts, auctions, increased funding access.
- Incentive **Strategies for AI Providers**:
 - Energy cost savings, compliance, reinvestment in sustainability.



Incentives for AI inference:
Balancing **accuracy vs. latency** via constrained optimization:
Through appropriate incentives, **carbon emissions can be reduced up to 80%** by accepting lower inference accuracy and **up to 25%** by accepting higher inference latency

THANK YOU FOR YOUR ATTENTION

