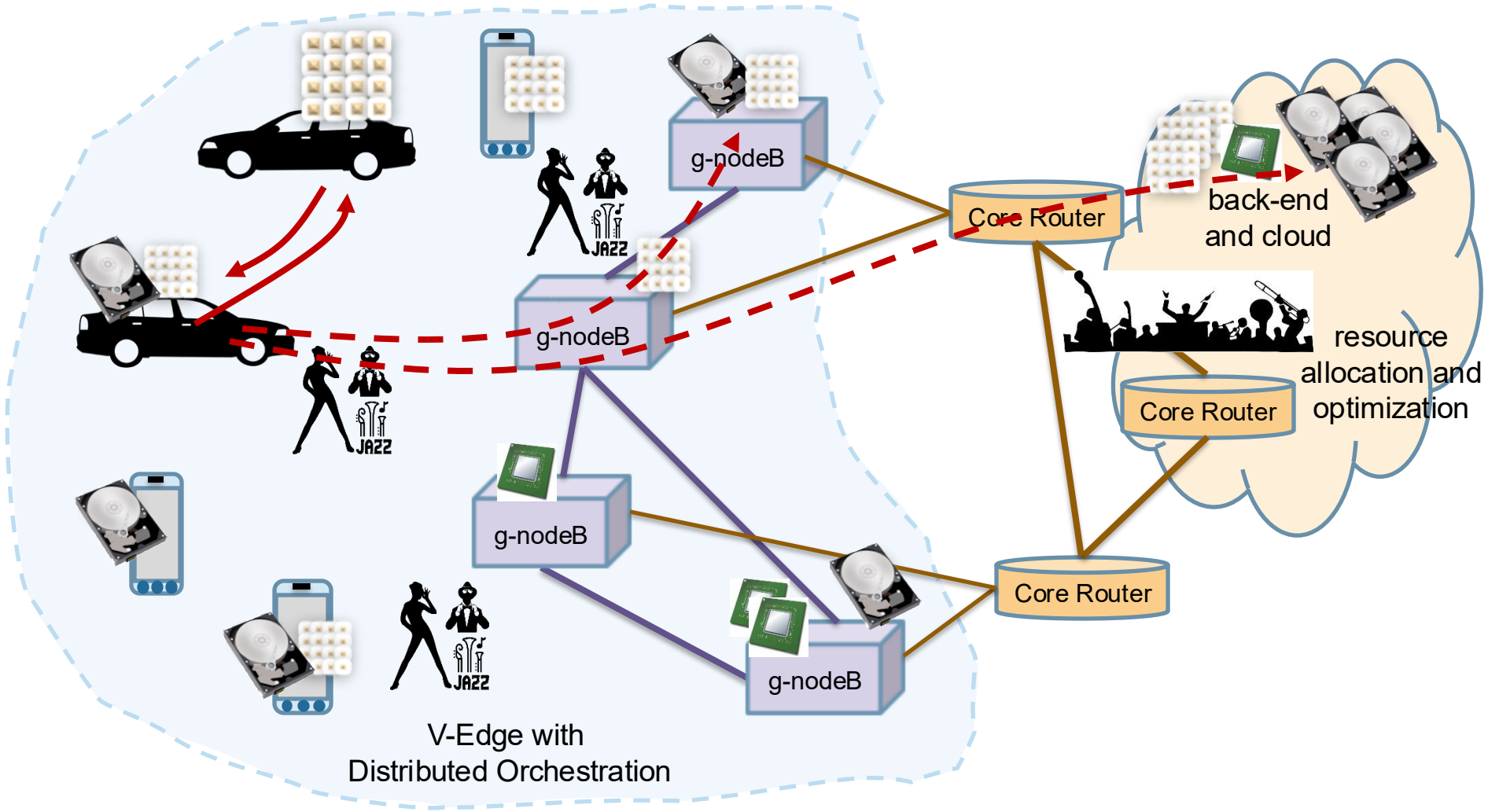
Technische Universität Berlin

Data-driven IoT Systems: Emerging Verticals, Challenges, and Opportunities – An Edge-based View

Falko Dressler

From Edge to V-Edge

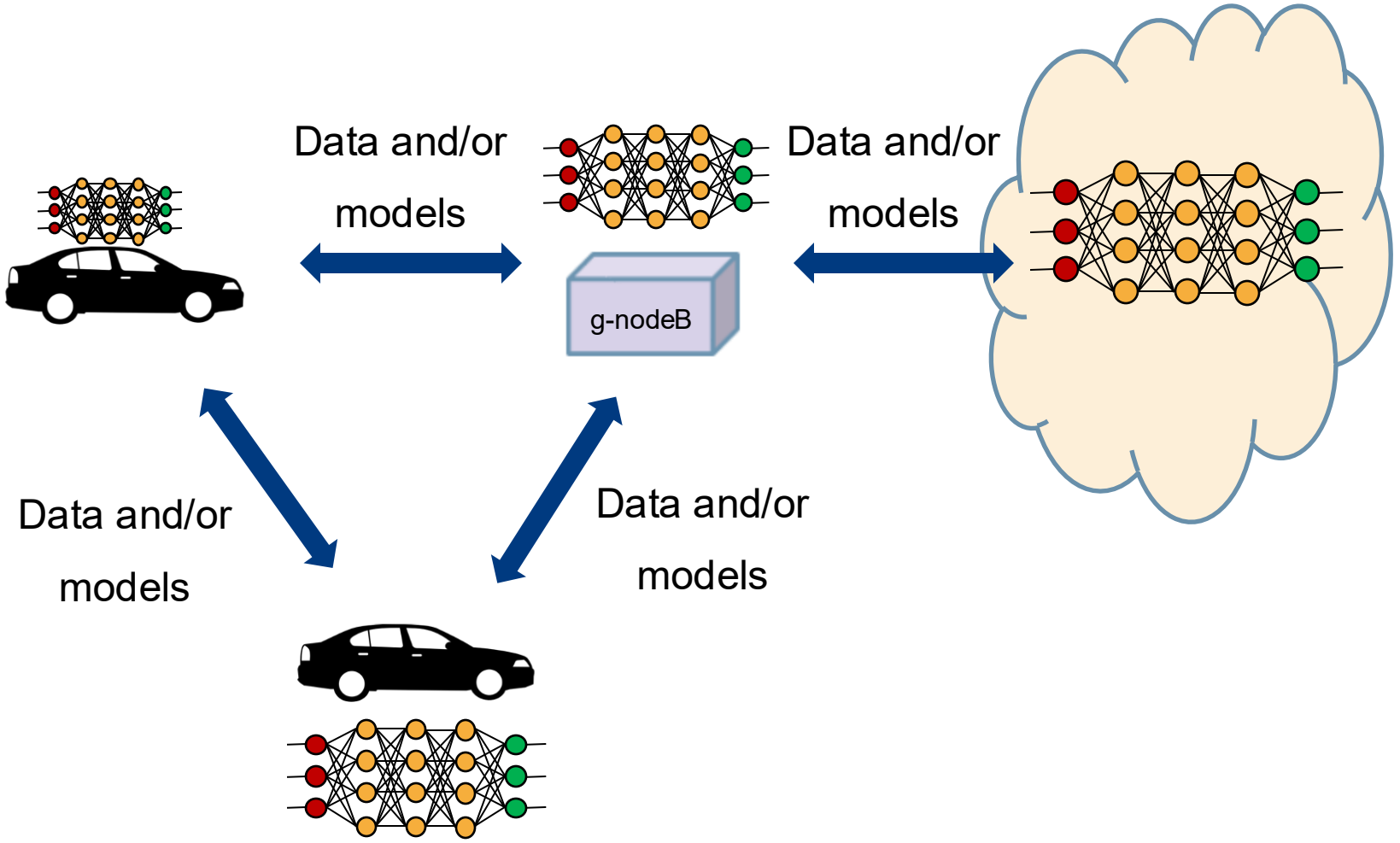
Virtualized Edge Computing



Falko Dressler, Carla Fabiana Chiasserini, Frank H. P. Fitzek, Holger Karl, Renato Lo Cigno, Antonio Capone, Claudio Ettore Casetti, Francesco Malandrino, Vincenzo Mancuso, Florian Klingler and Gianluca A. Rizzo, "**V-Edge: Virtual Edge Computing as an Enabler for Novel Microservices and Cooperative Computing**," IEEE Network, vol. 36 (3), pp. 24–31, May 2022.

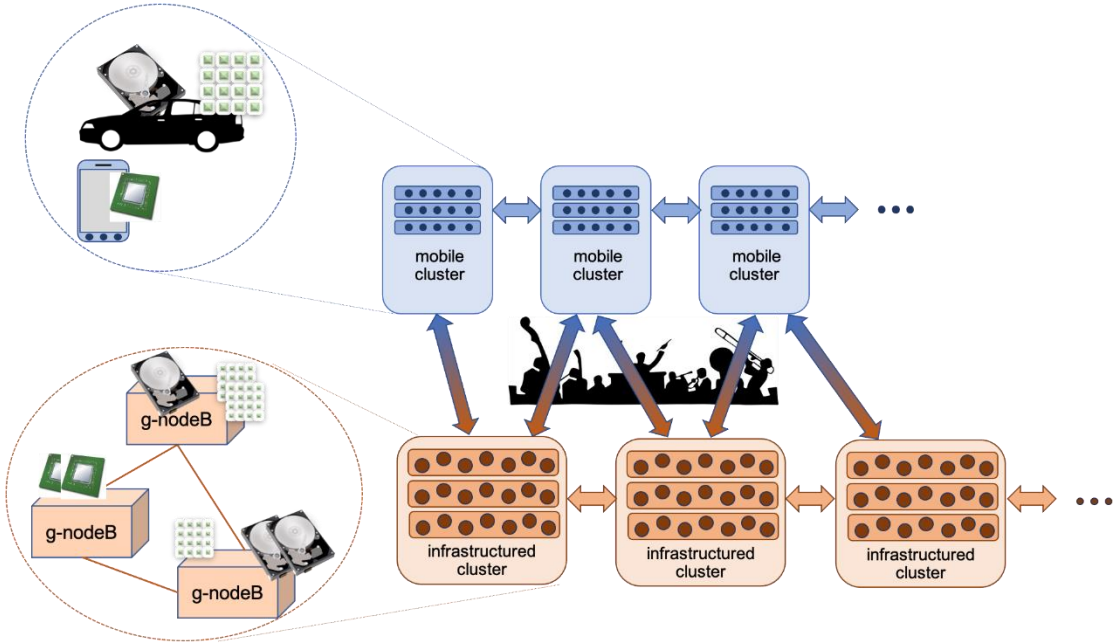
AI/ML and Virtualized Edge Computing

- Federated approaches to the extreme



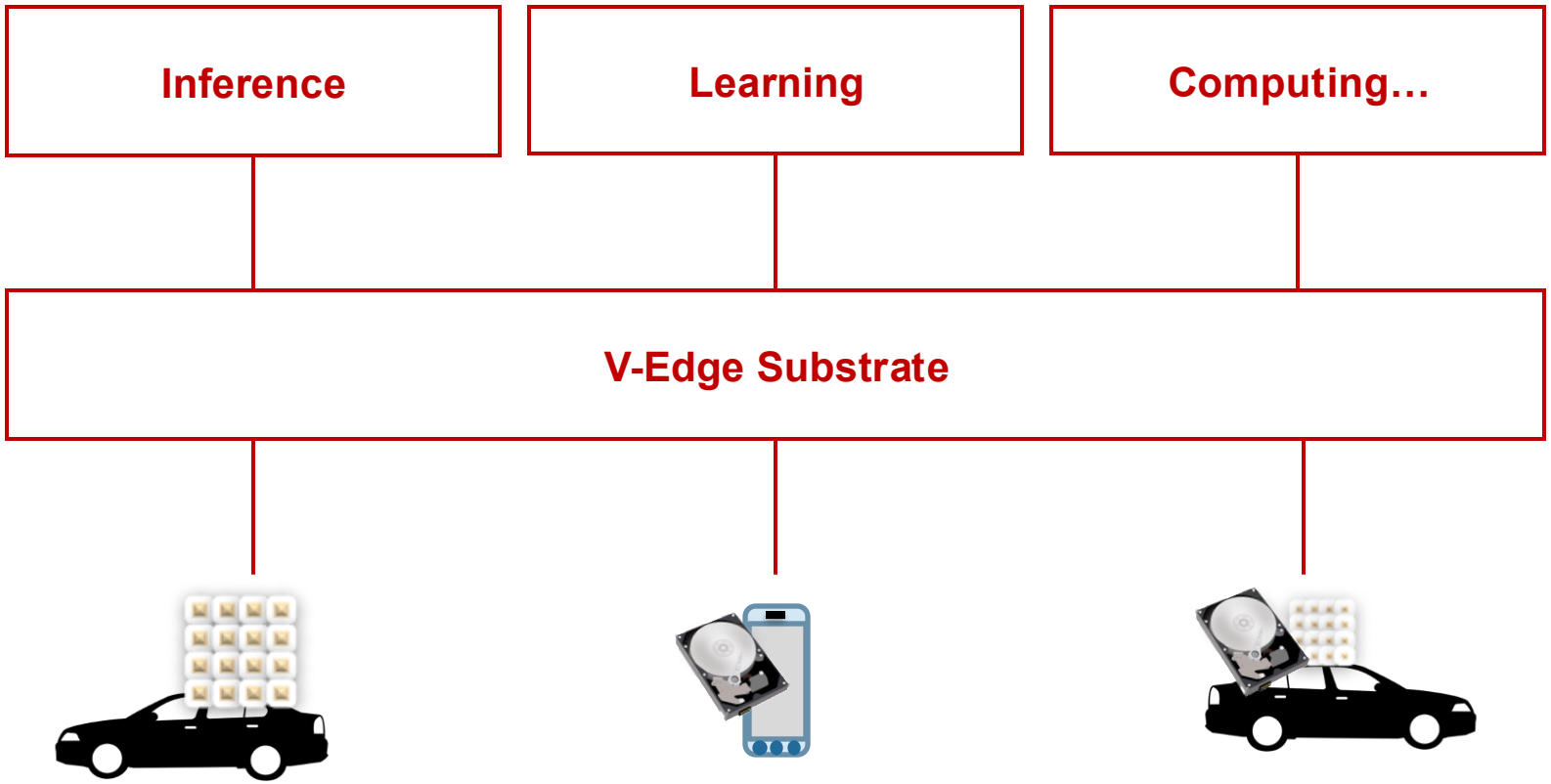
V-Edge Architecture

- **Logical resources** from mobile users and infrastructure-based systems (e.g., edge servers co-located with a gNB) are **aggregated into clusters**
- Multiple clusters are appropriately coordinated and **microservices can migrate from one cluster to another** to optimize the service location
- **Resource management** is done **by an orchestrator**, which may interwork with others, controlling neighboring clusters, **to migrate services**

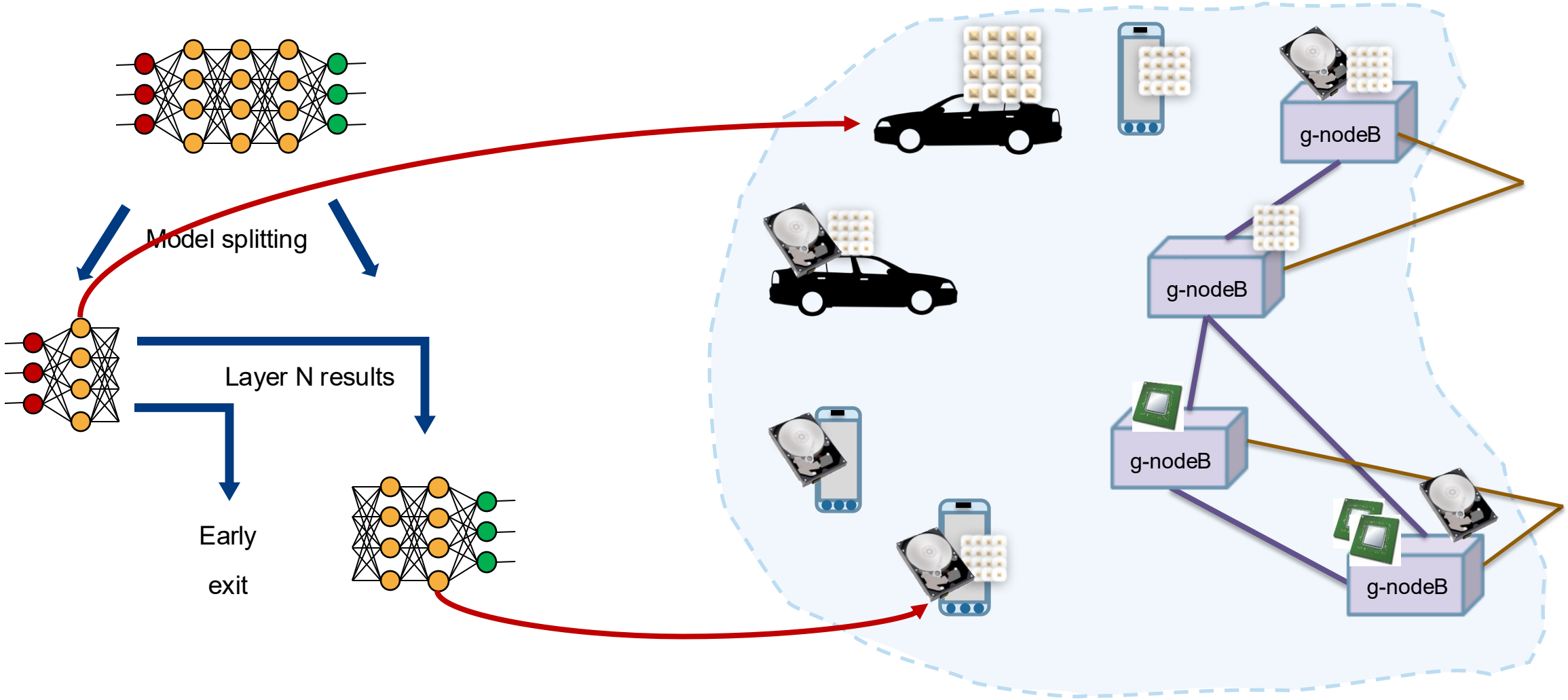


V-Edge Substrate for Distributed Computing at the Edge

Model Splitting: Distributed Inference, Distributed Learning, and more

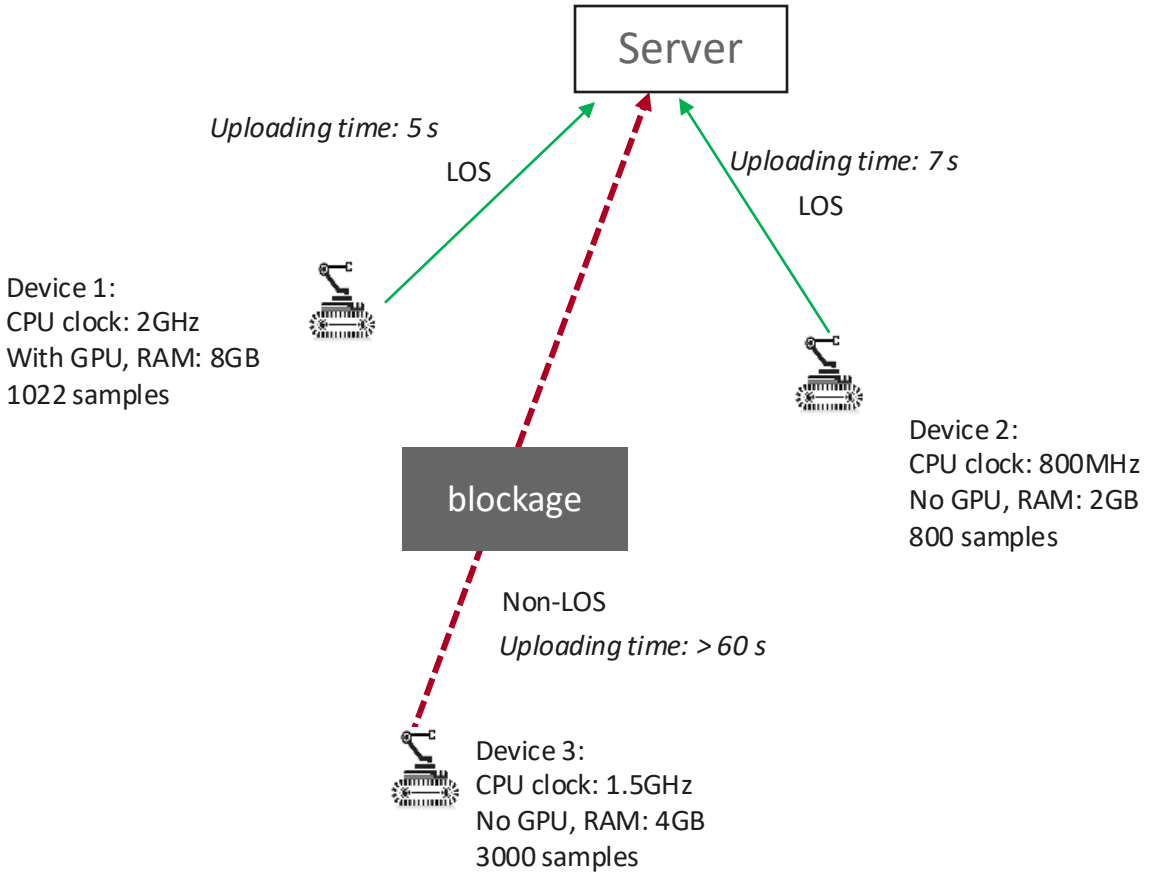


Model Splitting



Machine Learning within the V-Edge

Heterogeneity: the Motivation for FL Variations

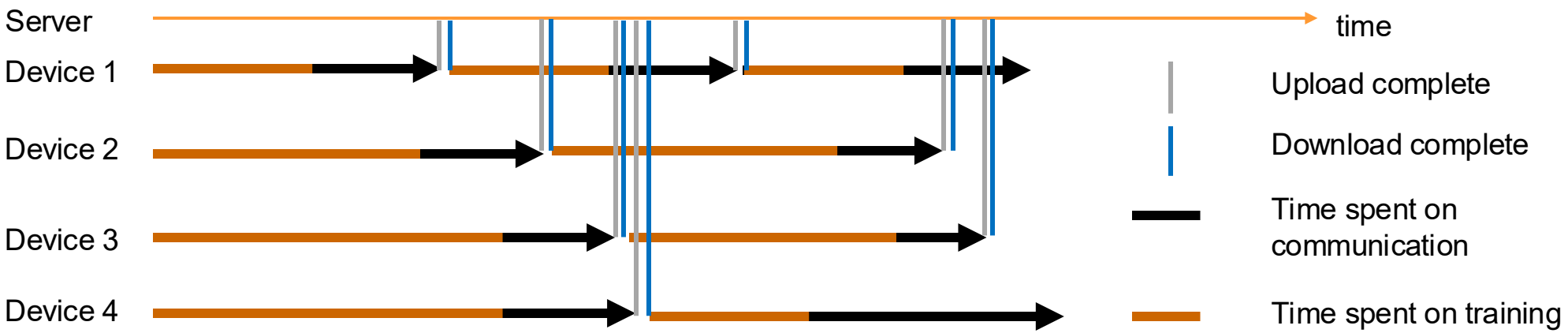


- Federated Learning (FL): with data distributed across edge devices, we aim to exploit their data and computational capacity.
- For devices to share their learned knowledge through models to the server.

- Classic FL involves strictly synchronized model exchange, which is difficult considering heterogeneity of:
  - Data distribution
  - Computational capacity
  - Communication link throughput

Asynchronous FL

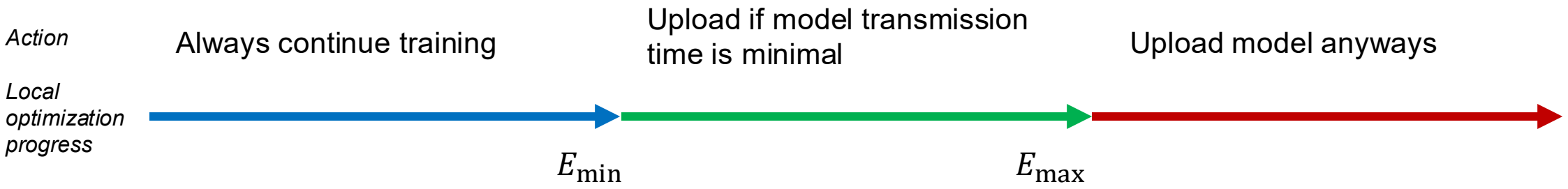
- By relaxing strict synchronization cycles, we allow FL devices of various capability to upload model upon completion of their training



Mengfan Wu, Mate Boban and Falko Dressler, "Flexible Training and Uploading Strategy for Asynchronous Federated Learning in Dynamic Environments," IEEE Transactions on Mobile Computing, vol. 23 (12), pp. 12907–12921, December 2024.

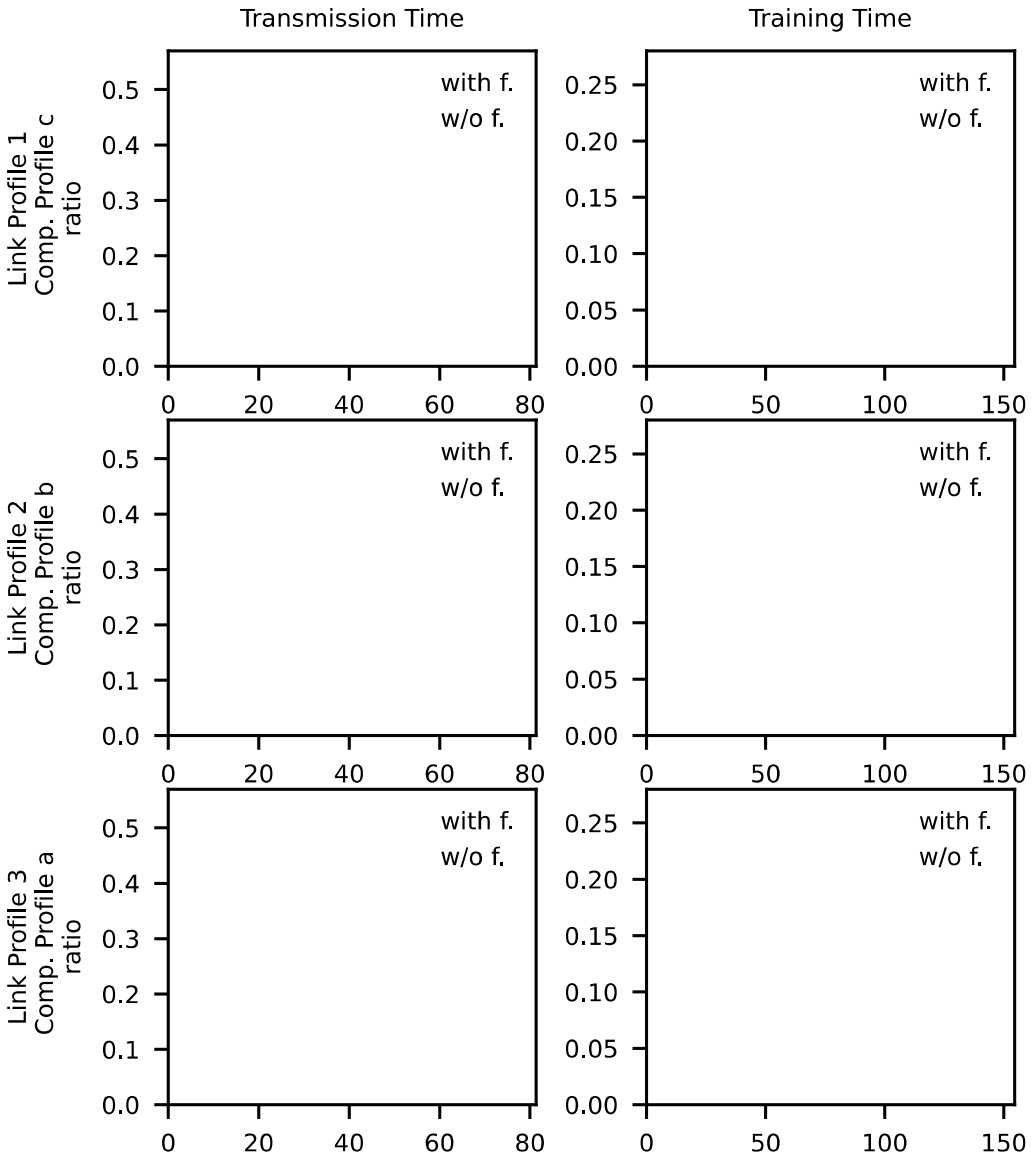
Asynchronous FL with Resource Maximization

- Assuming FL devices know/predict the throughput of their links to the server, we set a range for training process to be acceptable: $[E_{\min}, E_{max}]$
- Devices are aware of their training process and monitor the future link throughput within a time span



Training time vs. transmission time

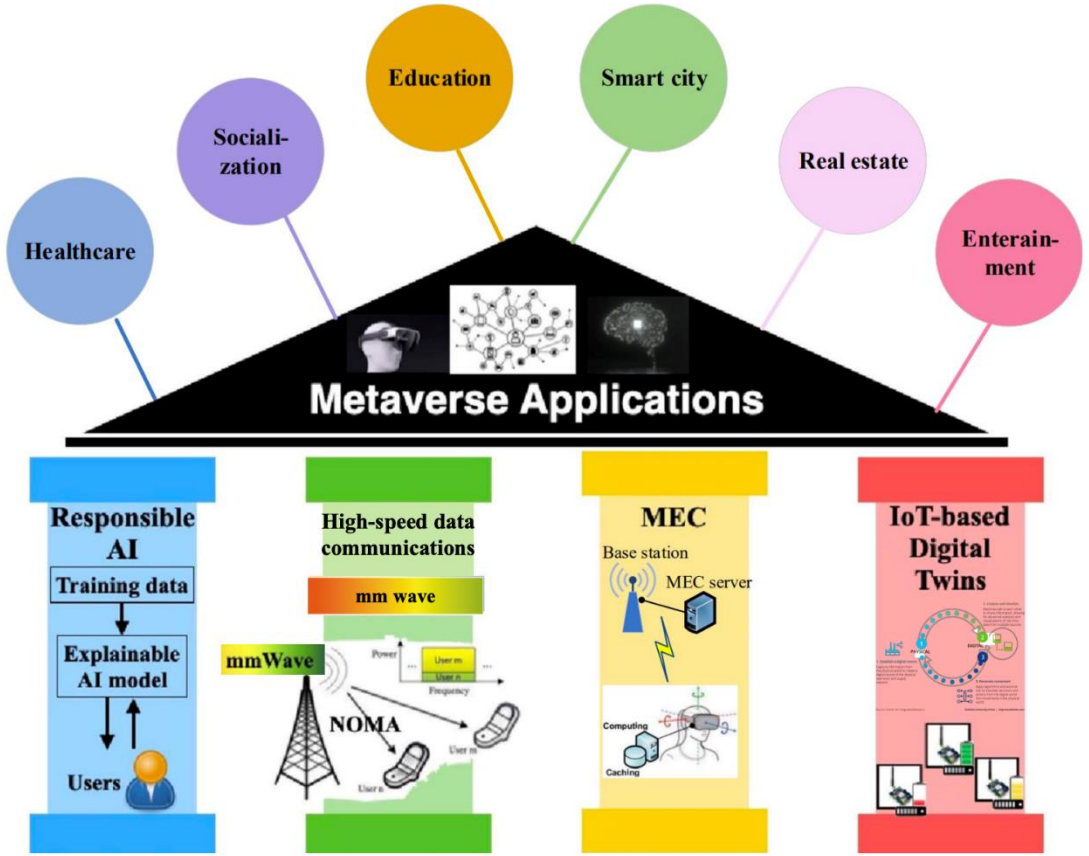
- Typical vehicular scenario



Pillar for future resilient Metaverse applications

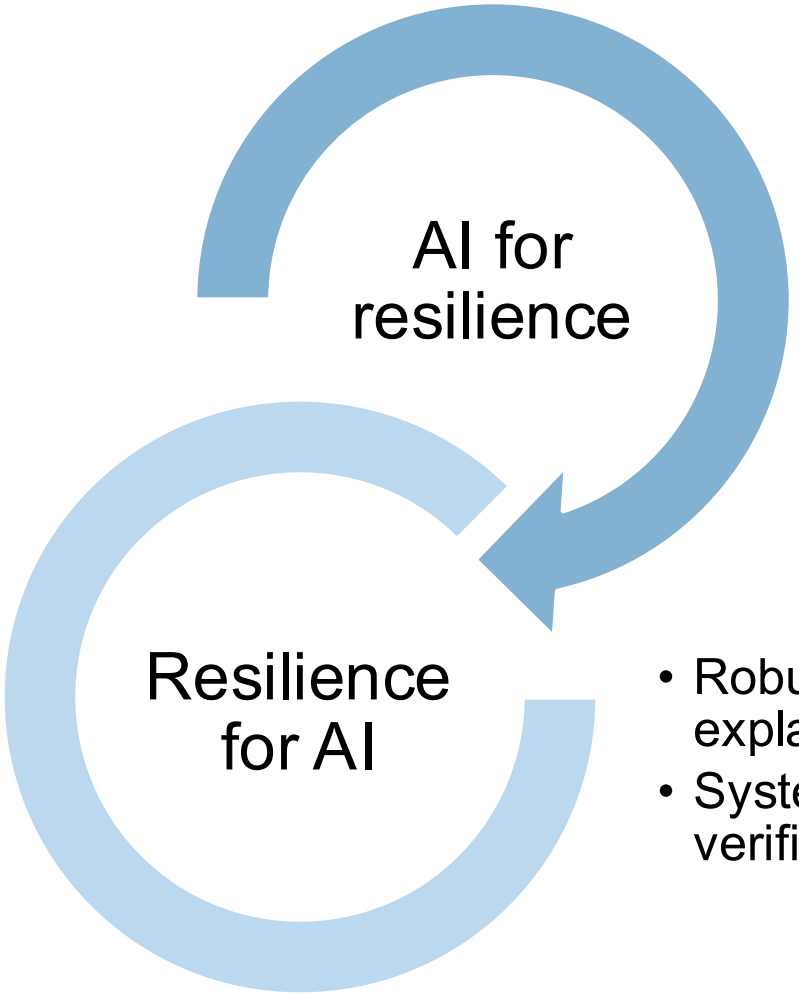
Task Offloading at the Edge – towards the MetaVerse

- Lots of AR/VR/XR but everything else is **AI/ML**
  - Training machine lear



Kai Li, Yingping Cui, Weicai Li, Tiejun Lv, Xin Yuan, Shenghong Li, Wei Ni, Meryem Simsek and Falko Dressler, "When Internet of Things meets Metaverse: Convergence of Physical and Cyber Worlds," IEEE Internet of Things Journal, vol. 10 (5), pp. 4148–4173, March 2023.

Edge Computing with Deeply Integrated Learning



- Self-learning and adaptive protection for the connected world
- Automatic detection, mitigation and remediation of threats

- Robust, secure, efficient, and explainable AI components
- Systematic analysis, testing, and verification of resilience in AI