Università degli Studi di Messina

NDN for AI and AI for NDN: challenges, opportunities and early results

Marica Amadeo[*], Claudia Campolo[§], Antonella Molinaro[§]

*University of Messina, Italy
§University Mediterranea of Reggio Calabria, Italy

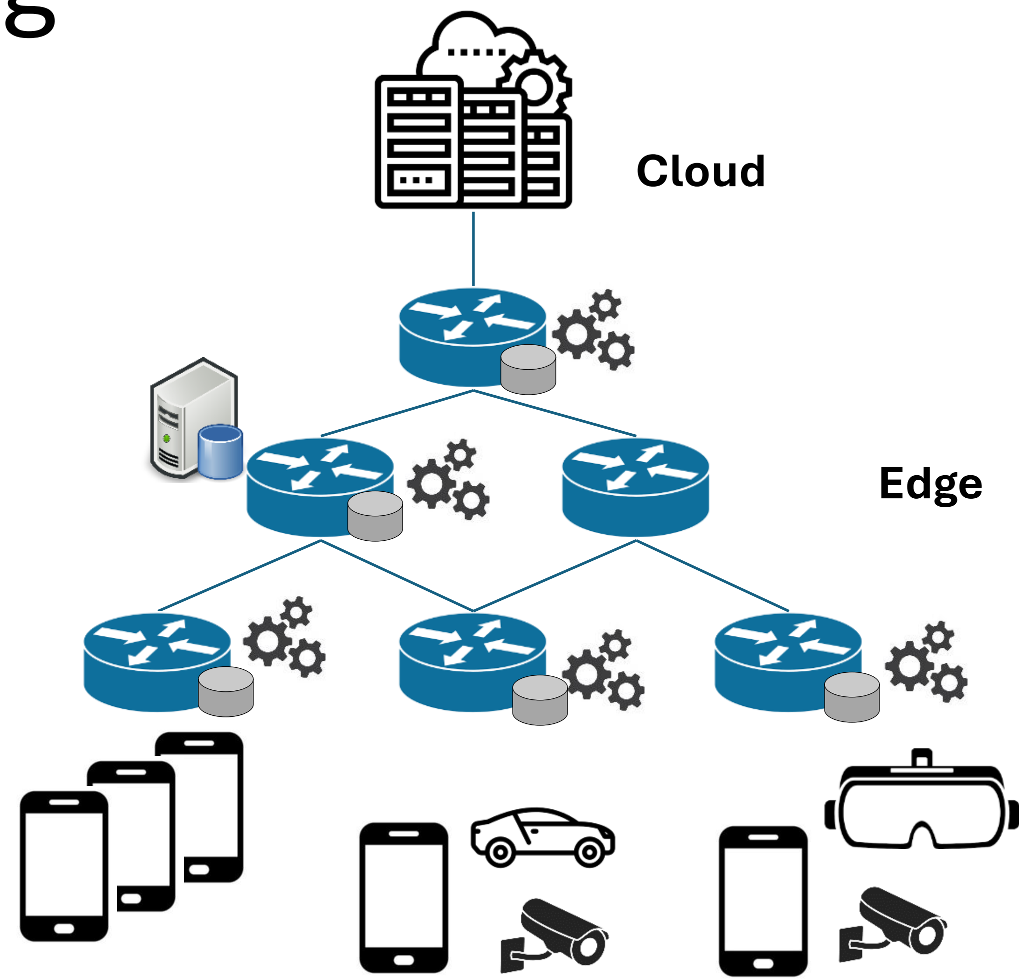
Outline

- **Two complementary perspectives on the NDN ↔ AI interplay**
- **NDN for AI**
  - AI applications are content-centric
  - NDN benefits
  - Early results: the case of Federated Learning and split inference
- **AI for NDN**
  - AI can improve NDN content discovery and delivery
  - Semantic-aware NDN forwarding
  - Early results: The case of vehicular crowdsensing

NDN for AI

Distributed AI and networking

- Distributed AI is moving at the network edge:
  - Federated Learning, Edge AI, Split inference
- AI applications increasingly require networking support beyond packet forwarding
  - <u>In-network computing</u>
- The network is now a bottleneck
  - The current Internet remains host-centric



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm’26

AI applications are content-centric

- Distributed AI exchanges:
  - models, embeddings, inference results…
- These are naturally named data objects
- AI workflows are iterative, content-centric, and cache-friendly

Yet, the network is still host-centric!

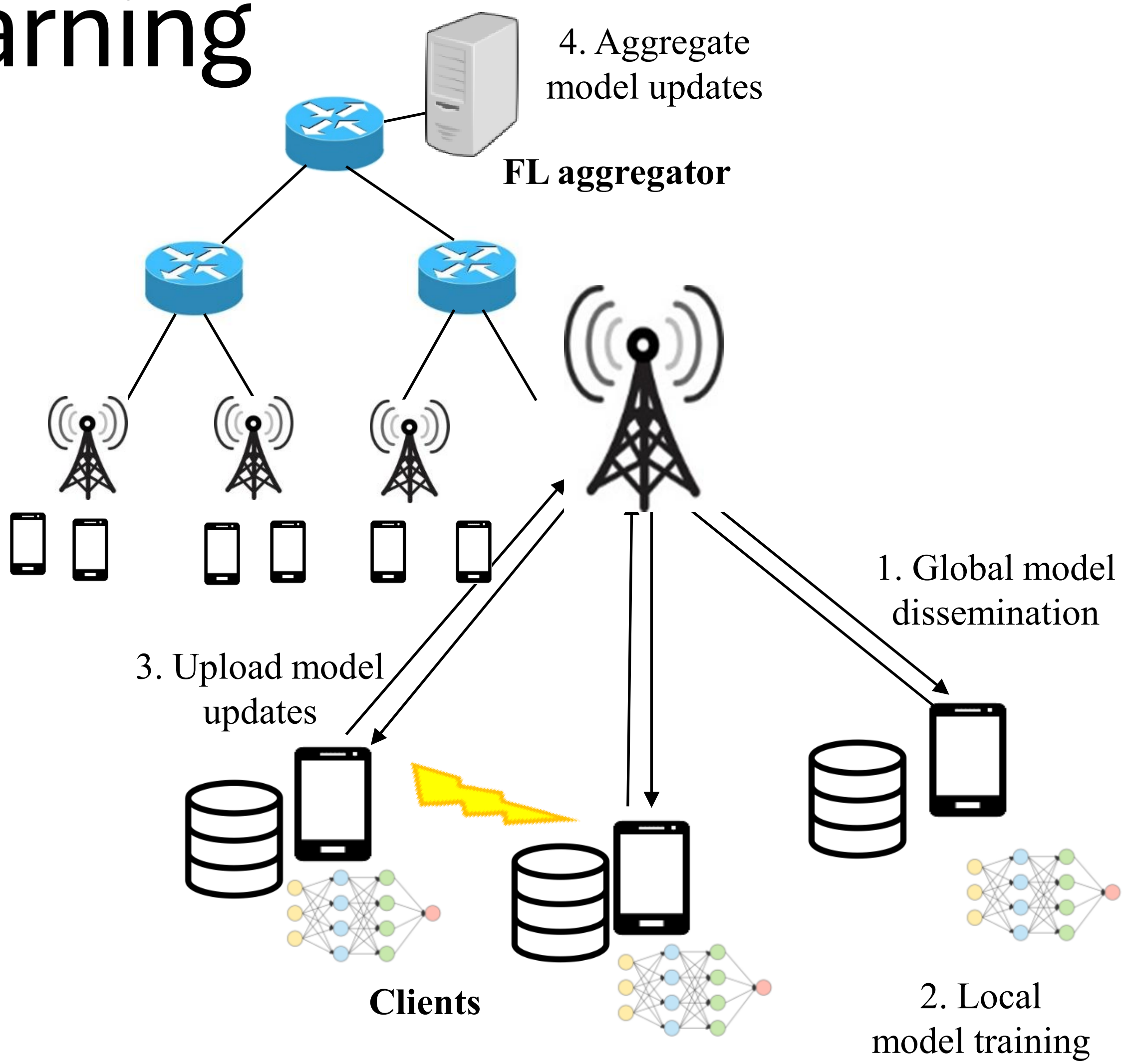
NDN as an enabler of distributed AI

NDN Feature	Benefits
Name-based delivery	AI tasks, models, inputs, and inference results can be identified by name, independently of the host. In-network processing is facilitated
In-network caching	Models and results can be served by nearby caching nodes, reducing latency and increasing resiliency to losses
Multicast support	Efficient distribution of models and results to many clients at the same time
Stateful forwarding	Fast recovery in presence of packet losses
Security by design	Integrity, authentication, and provenance of models, and inference results

M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm’26

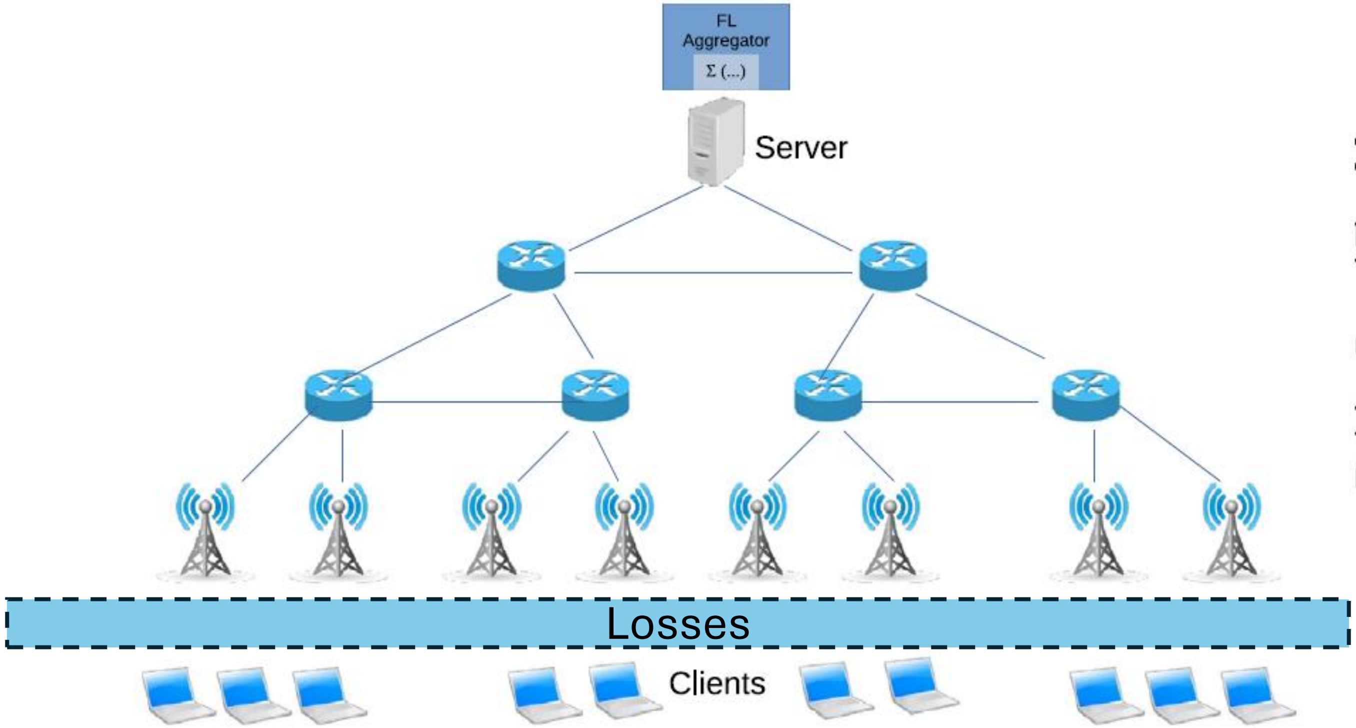
The case of Federated Learning

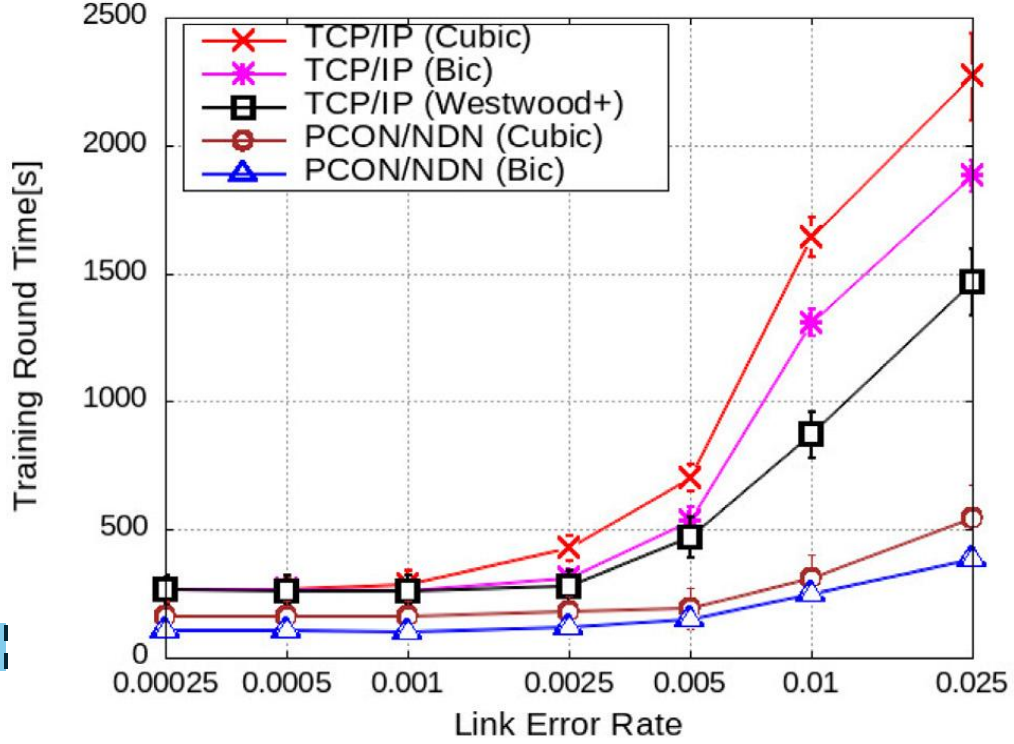
- FL systems assume:
    - end-to-end communication
    - no in-network computing
- Convergence depends on the slowest client (straggler issue)
    - Lossy wireless links increase round duration
    - End-host retransmissions preserve reliability but add delay
    - Excluding stragglers hurts model accuracy



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm’26

Improving FL via NDN



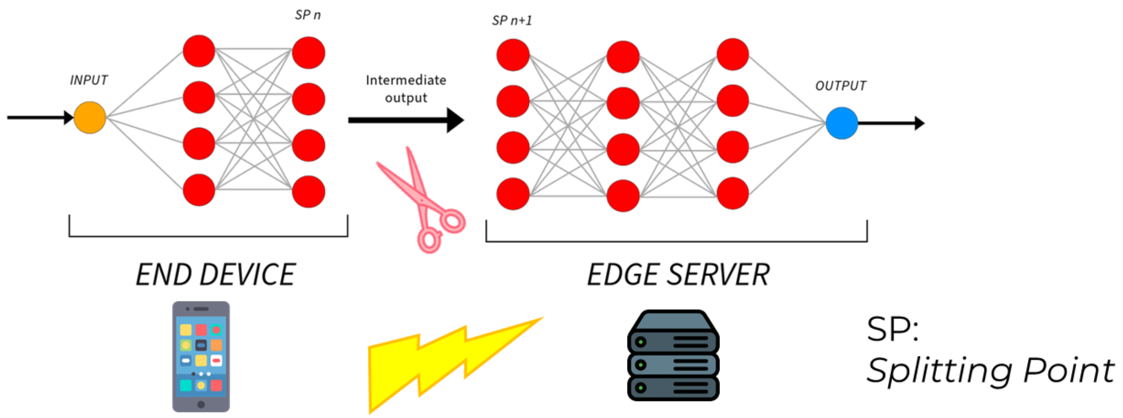


PCON (Practical CONgestion control) + NDN
vs TCP/IP

M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

Improving Split Inference with NDN

- Split Inference splits Deep Neural Networks (DNN) across multiple nodes to improve efficiency and reduce the burden on individual devices
  - The inference task is partitioned across multiple in-network nodes
- Input data, intermediate outputs and inference results to be transferred quickly and reliably

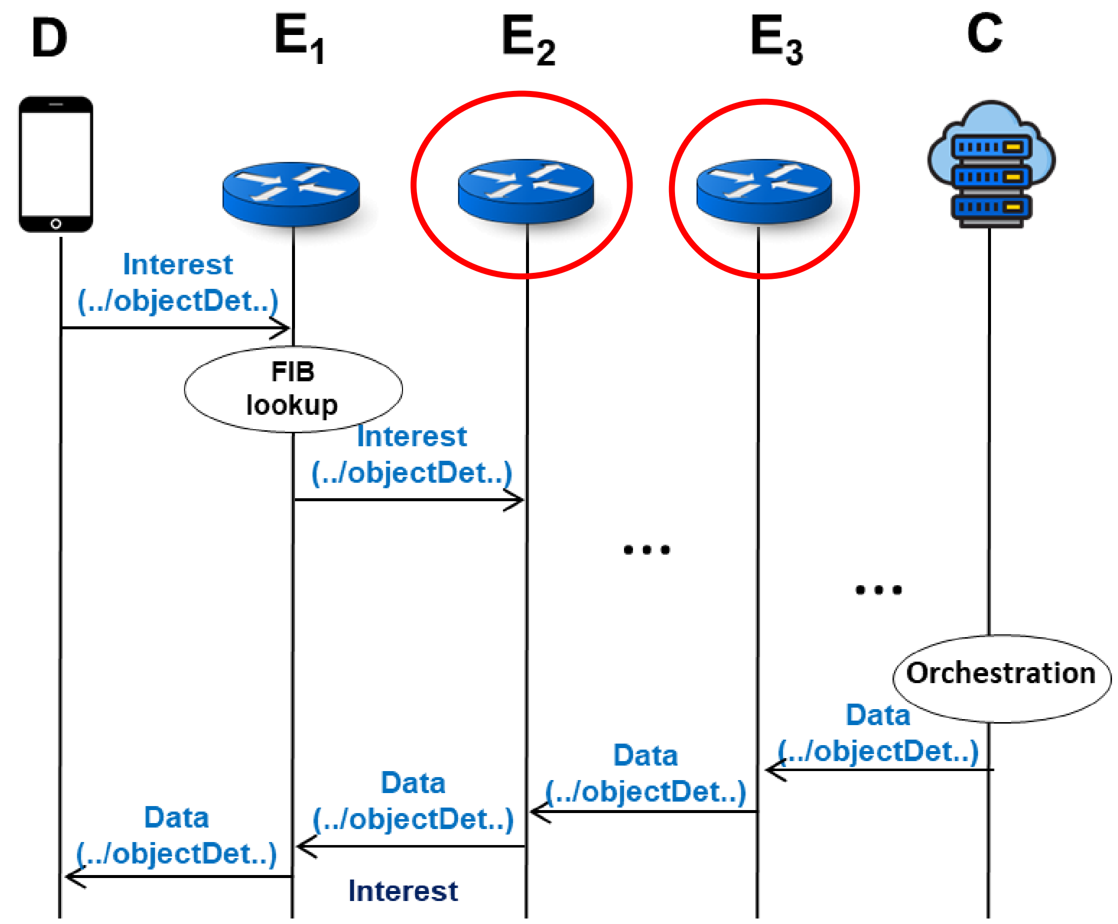


M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

Split inference via NDN

Discovery stage:

- Clients issue an inference request through an augmented Interest
- On-path edge nodes look up in the FIB and can candidate as executors
- The cloud decides the splitting strategy and instructs selected nodes

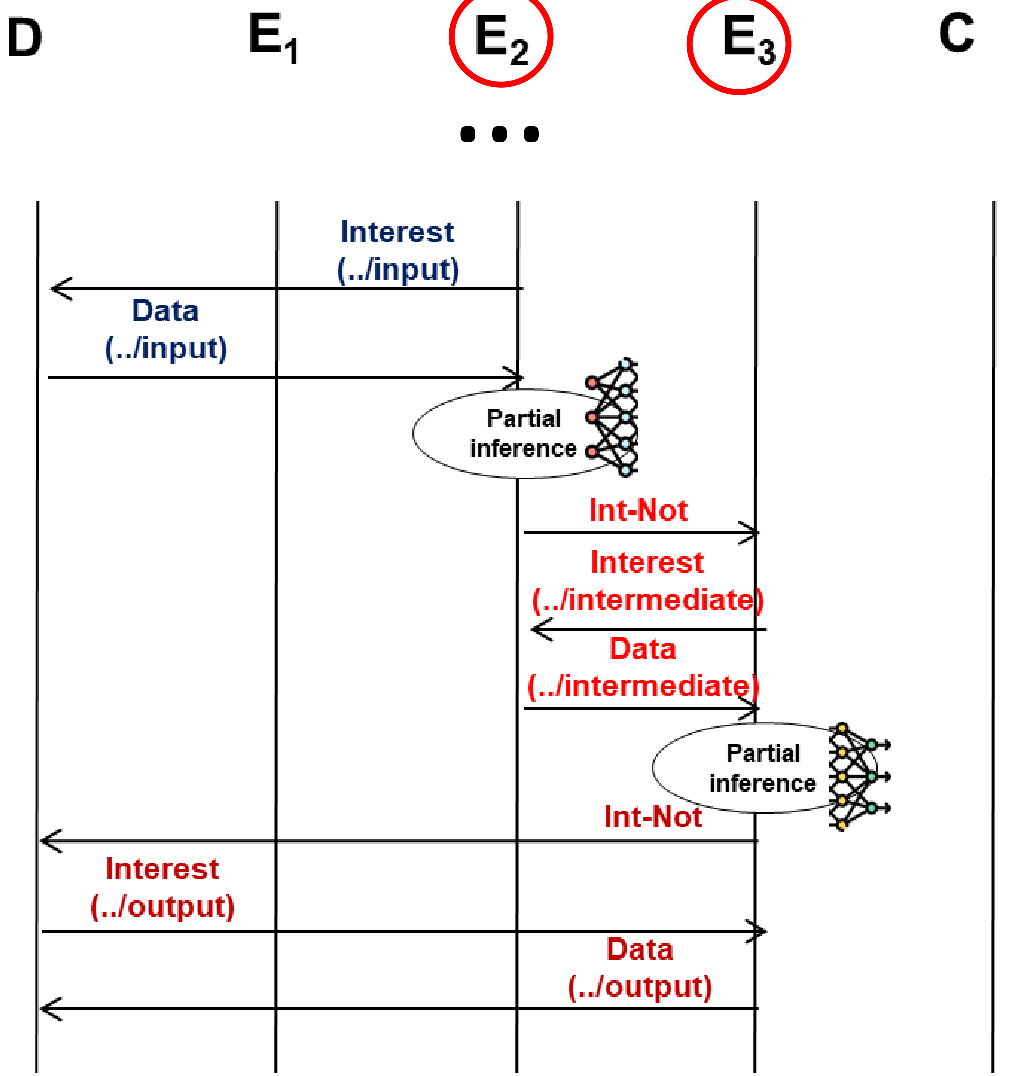


INTEREST
(/splitAI/objectDetection/ResNet50/img123.jpg/acc=90/lat=0.5)

...

M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

Split inference via NDN

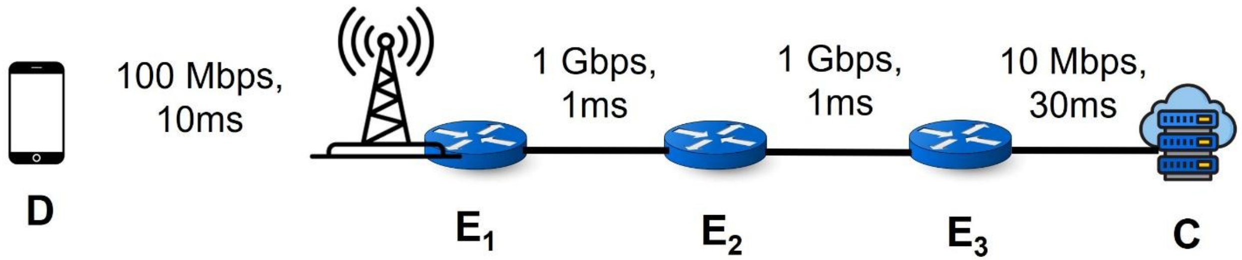


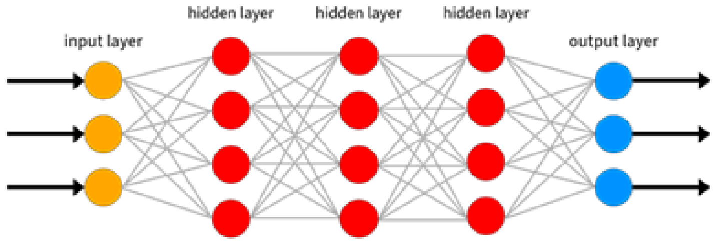
Data exchange

- Selected nodes retrieve the (input/intermediate output) data upon which executing their part of the DNN
- Upon completion, each selected node issues an Interest Notification (*Int-Not*) message to signal that the intermediate output/inference results is ready
- Data can be lost
  - Thanks to native NDN in-network caching intermediate nodes can help recover losses

M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

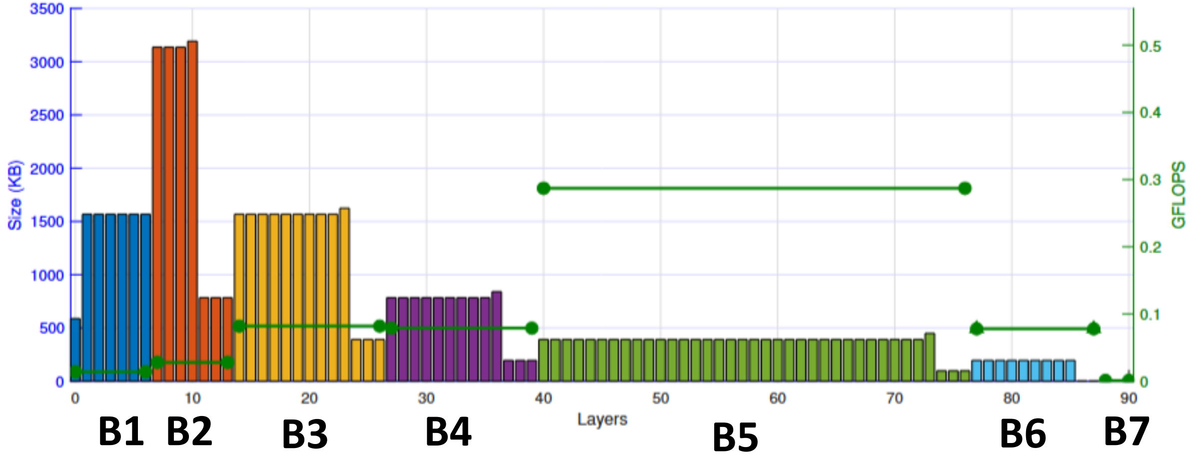
Split inference via NDN





ns-3 NETWORK SIMULATOR + ndnSIM

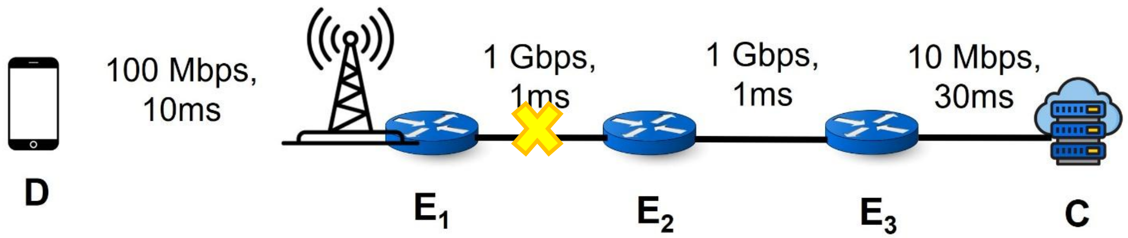
- Chain topology with 5 nodes
- DNN model: MobileNetV1 divided into 7 consecutive blocks
- Input data size: 330 kB



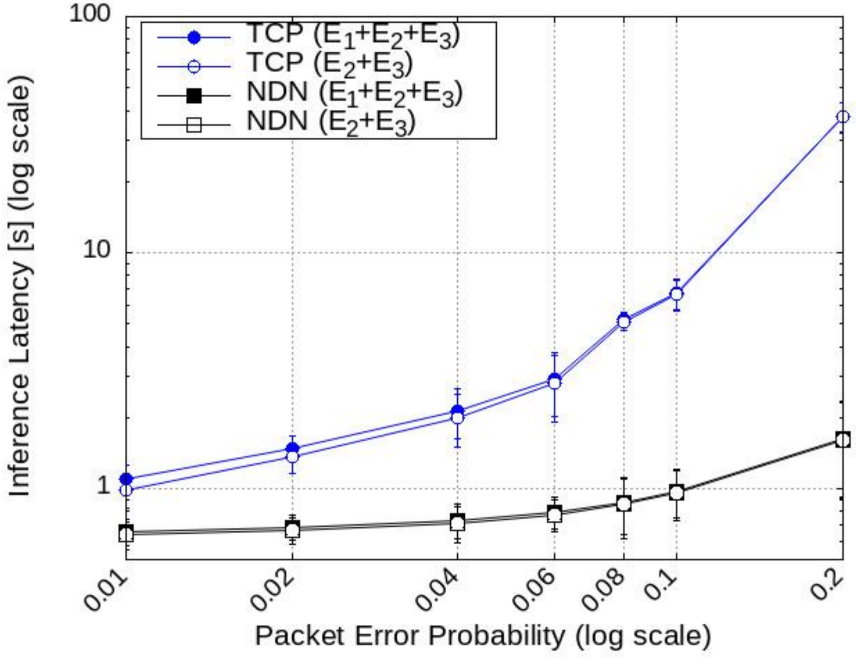
M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

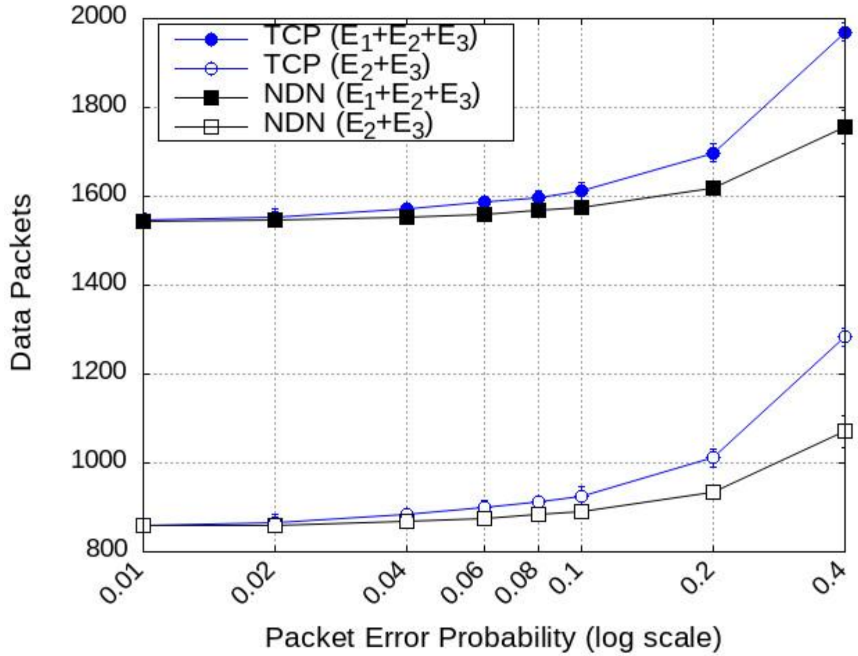
Split inference via NDN

TCP VS NDN



- Distinct split strategies
- A client asks for a DNN task
- Input data size: 330 kB



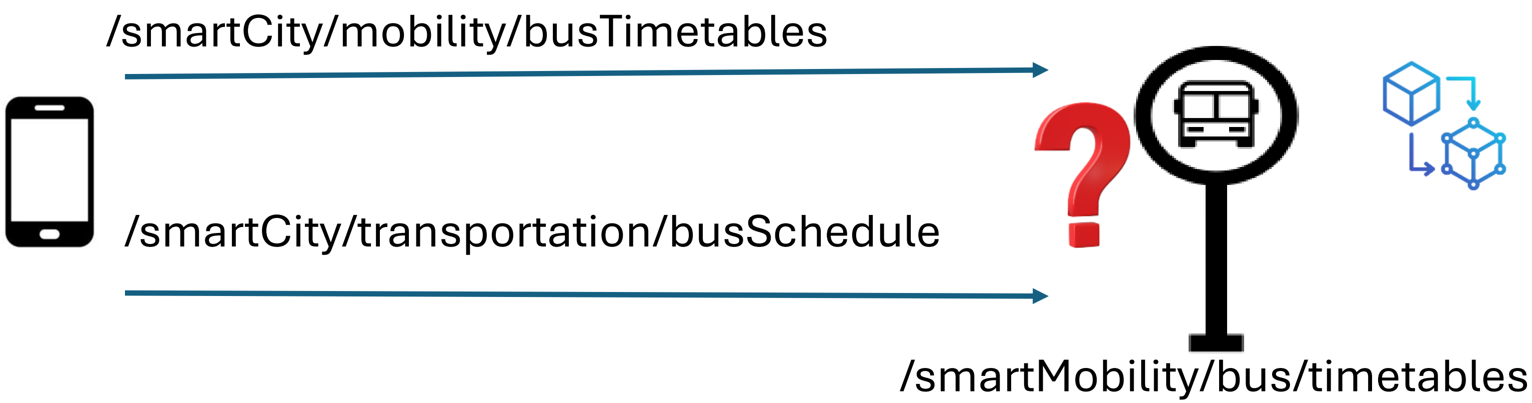


M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm’26

AI for NDN

Limitations of name-based lookups

- NDN relies on exact name match lookups
- But naming conventions may vary across data producers
- **Semantically equivalent data can have different names!**
- **We need a flexible data discovery solution**

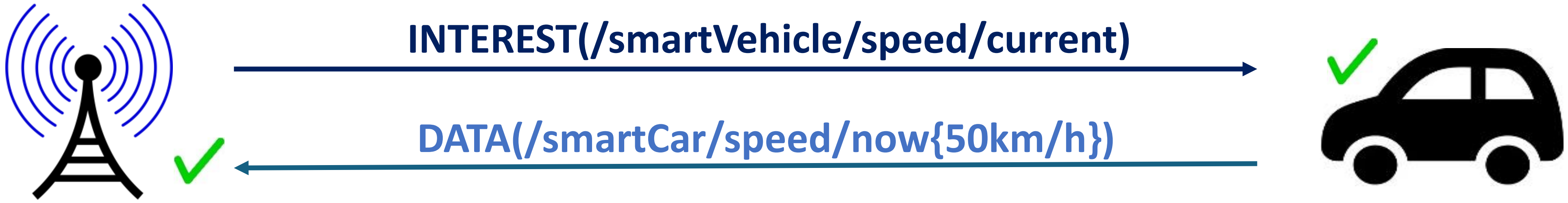


Idea:
semantic awareness in the NDN forwarding fabric

M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm’26

Augmenting NDN with semantic intelligence

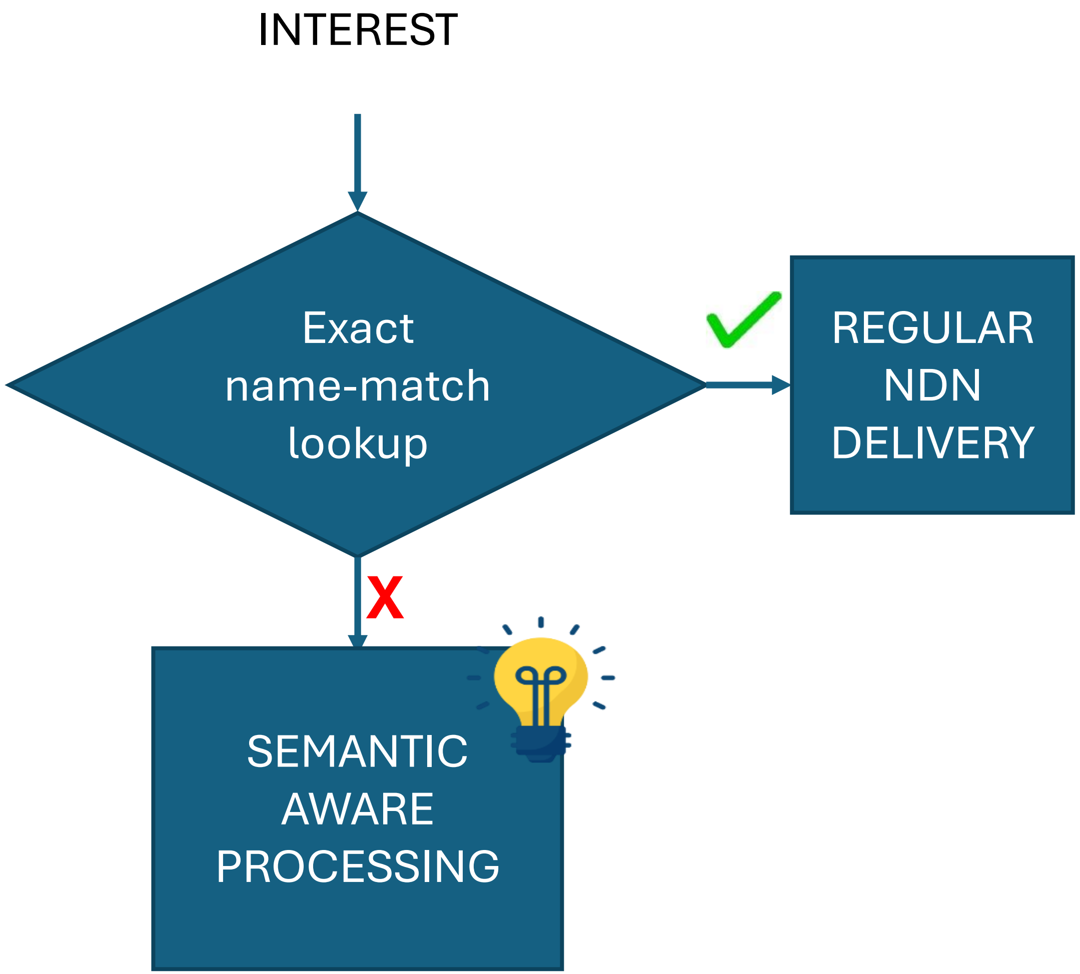
- A mechanism allowing NDN nodes to understand name semantics
- Implementing flexible name matching among different hierarchical names
  - We can use Natural Language Processing (NLP) mechanisms!
- First work: K. Chan, B. Ko, S. Mastorakis, A. Afanasyev, and L. Zhang, "Fuzzy interest forwarding," in Proc. 13th Asian Int. Eng. Conf., **2017**, pp. 31–37
  - single-component name matching, using Word2Vec to generate vector-based representations of name components and applying cosine similarity to measure their semantic closeness.



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

Augmenting NDN with semantic intelligence

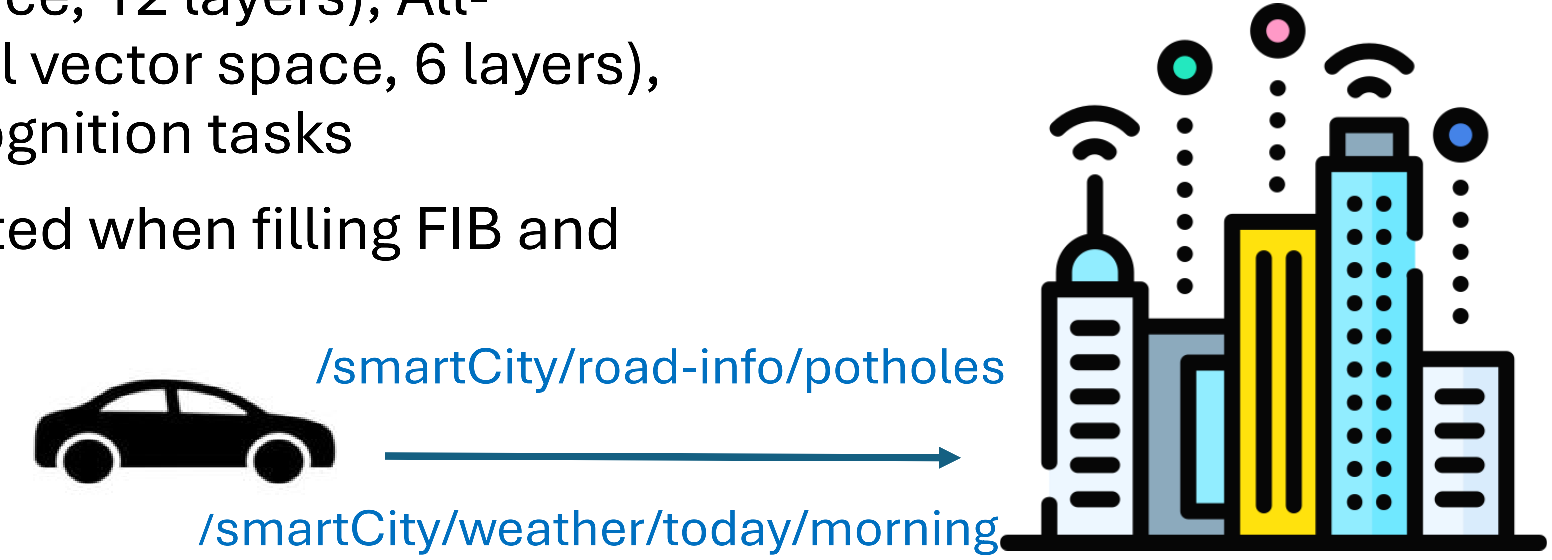
- We embed deep learning-based semantic recognition models into the NDN forwarding fabric.
- When an incoming data request fails to be solved with exact name matching, our approach looks for the **semantic similarity between the name in the Interest and the name of available data in the NDN tables**.
- This allows to discover and access data that are semantically similar, even when naming conventions differ.



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

Discovering smart-city services

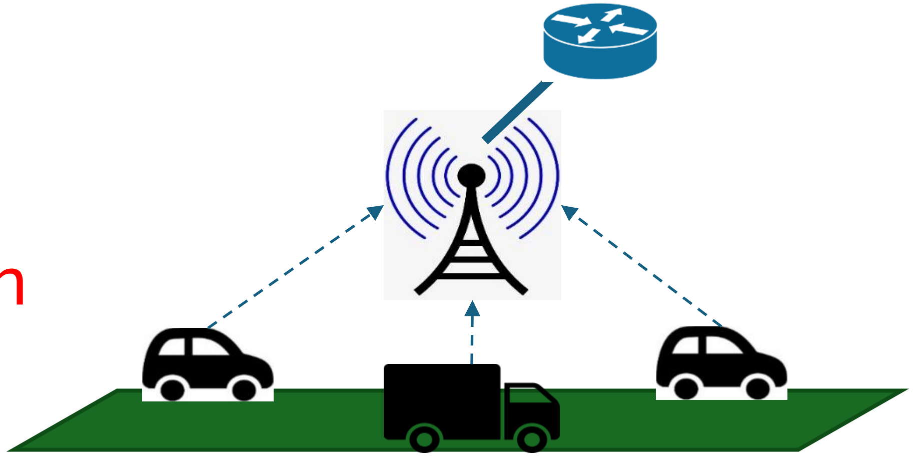
- Names are converted into embeddings (i.e., vectors) and compared in a multi-dimensional space
- Embeddings generation is based on pre-trained **Sentence Transformers models**, e.g., All-miniLM-L12-v2 (348-dimensional vector space, 12 layers), All-miniLM-L6-v2 (348-dimensional vector space, 6 layers), designed for fast semantic recognition tasks
- Name embeddings pre-computed when filling FIB and CS at the producer side



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm’26

The case of vehicular crowdsensing

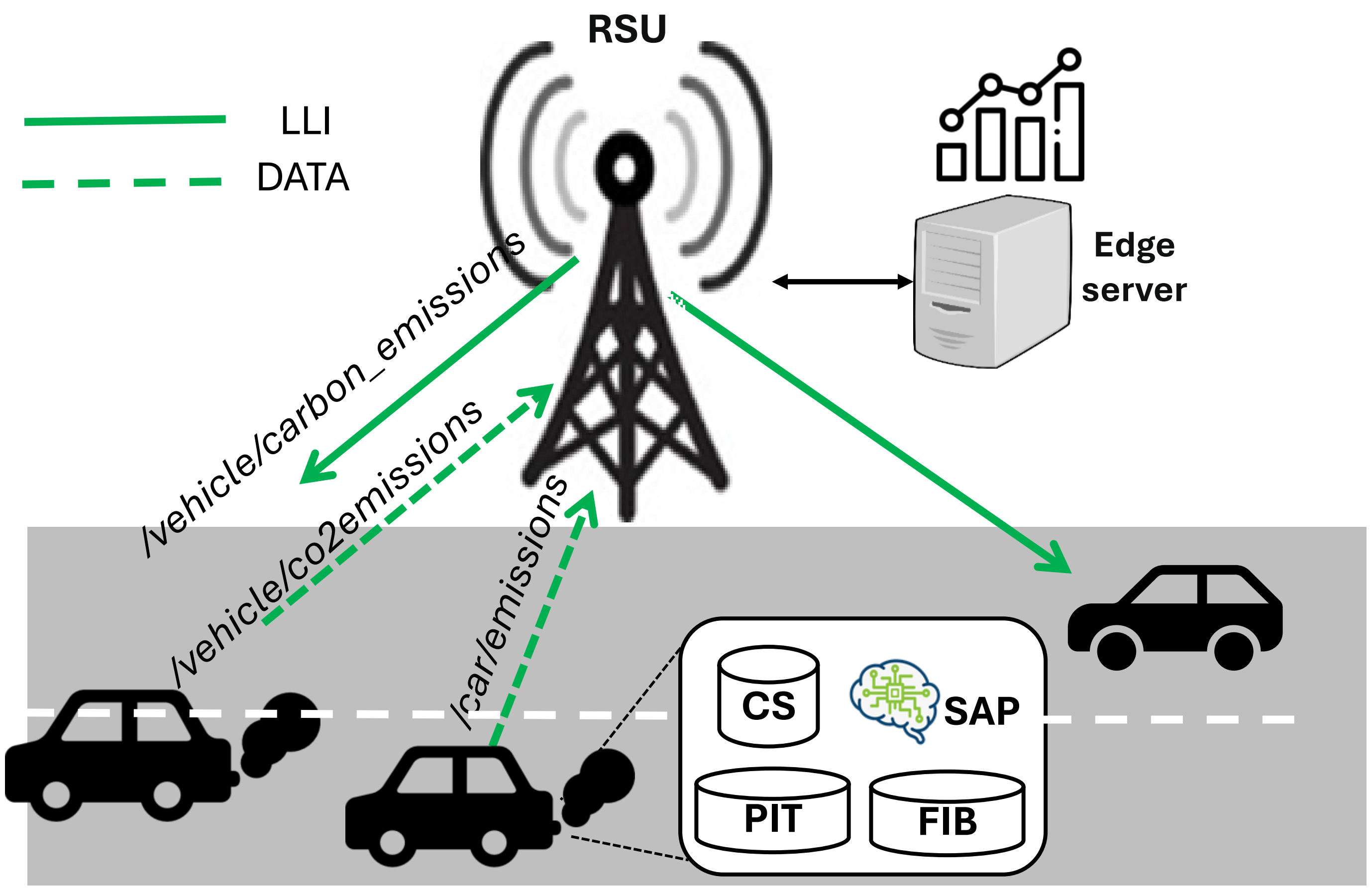
- Collecting named data with NDN:
  - ✓Vehicles' identity a priori unknown
  - ✓Handling mobility
  - ✓Data can be retrieved from vehicles with different manufacturers, different home countries -> Naming conventions may vary



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

The case of vehicular crowdsensing

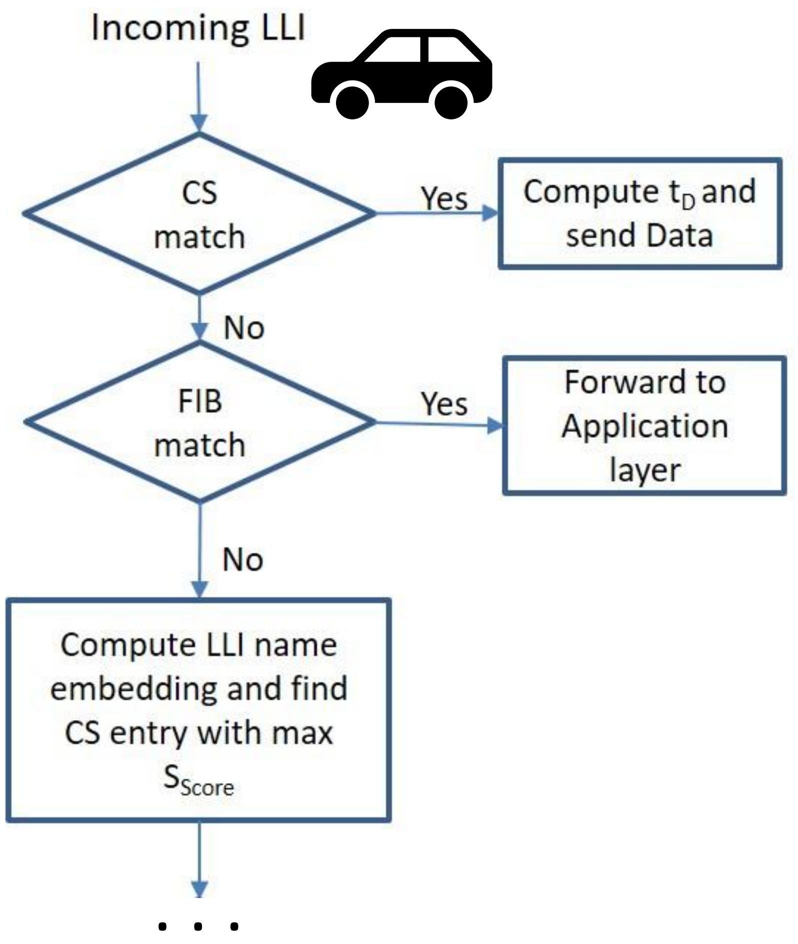
- Long Lived Interests (LLIs) are broadcasted to collect data from vehicles acting as mobile producers
  - ✓ **Semantic Aware Processing (SAP) is implemented!**
- Single-hop interactions: LLI propagation beyond RSU coverage is not considered



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

The case of vehicular crowdsensing

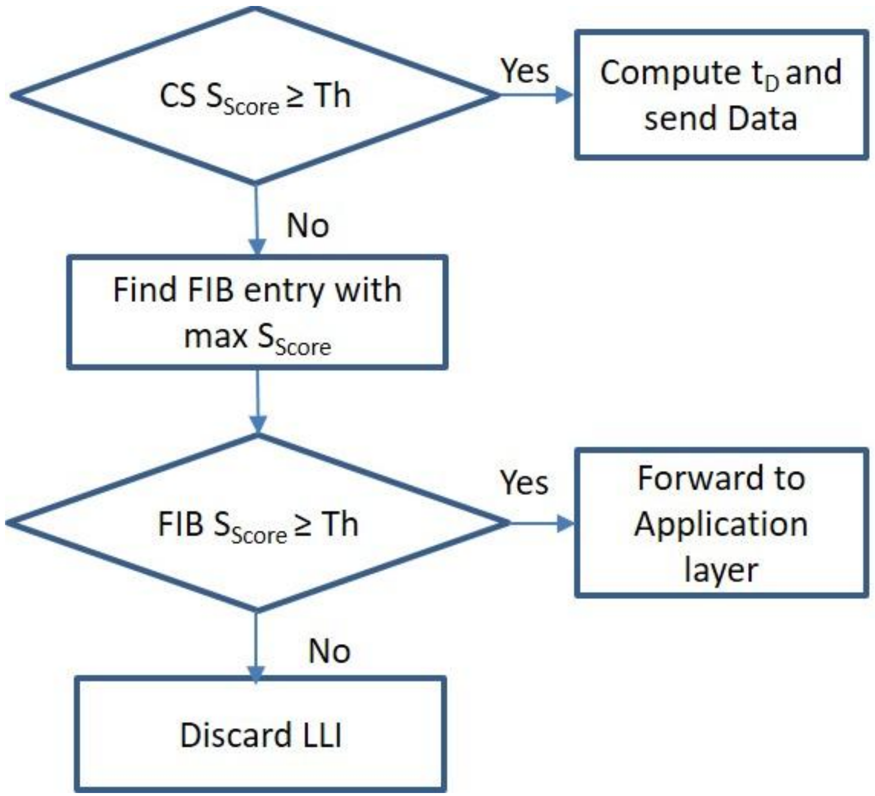
- After receiving the LLI, the vehicle looks for an exact name match in the CS or the FIB
- Data is transmitted after a random defer time to avoid collisions
- If no exact match is found, SAP is triggered
  - ✓ Embedding of the LLI name is compared against the embeddings of Data names
  - ✓ A semantic score in [0;1] is computed using the cosine similarity metric



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

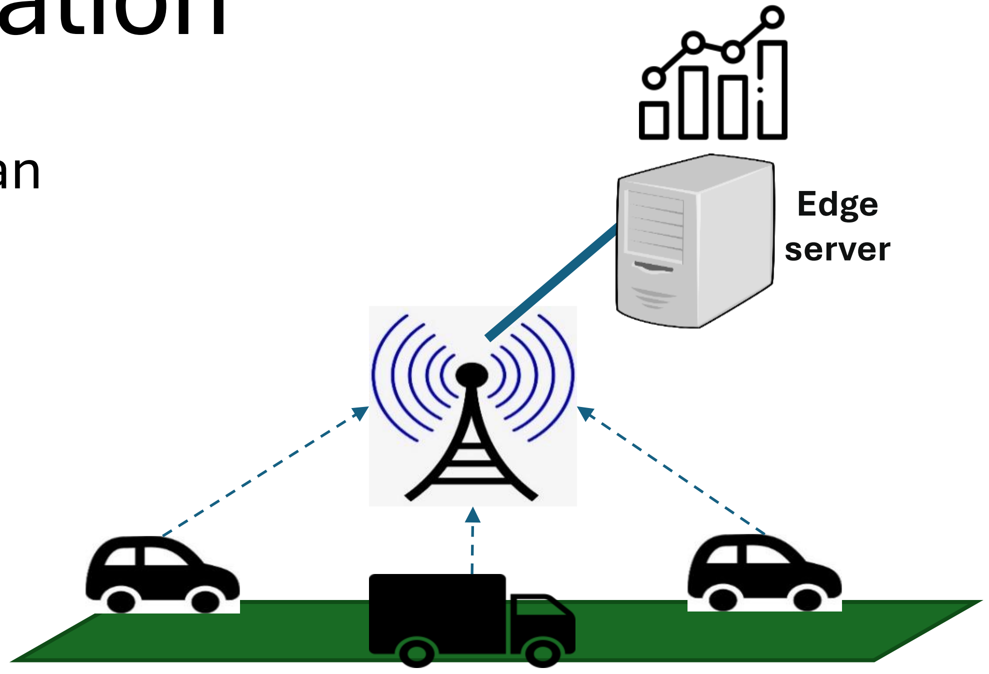
The case of vehicular crowdsensing

- Data with the highest semantic score, that is also higher than a similarity threshold, is scheduled for transmission
- If no semantic match is found in the CS, similarity with FIB names is considered
- If also semantic FIB matching fails, LLI is discarded



Semantic-aware NDN: Evaluation

- A crowdsensing service within the coverage area of an RSU
- A catalog of 500 LLI names
- Each vehicle can provide up to 100 contents
  - ✓ 5% exactly match LLI names
  - ✓ 95% have varying degree of similarity, including none
- SAP implemented with pre-trained Sentence Transformers models
  - ✓ All-mpnet-base-v2 (768-dimensional vector space)
  - ✓ All-miniLM-L12-v2 (348-dimensional vector space, 12 layers)
  - ✓ All-miniLM-L6-v2 (348-dimensional vector space, 6 layers)



M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm’26

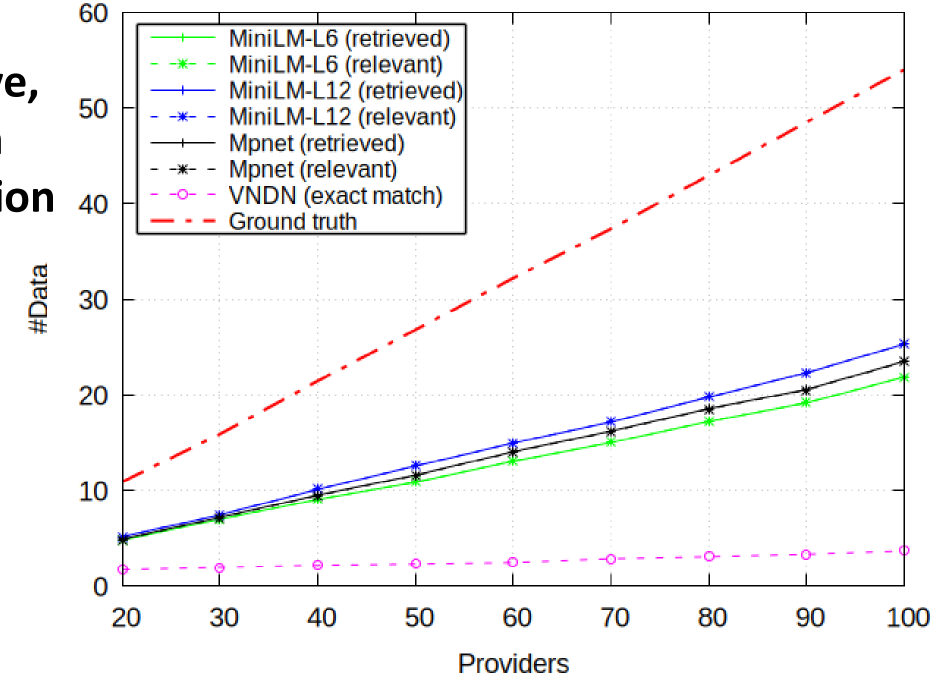
Semantic-aware NDN: Evaluation

AVERAGE COMPUTATION TIME PER INTEREST NAME FOR DIFFERENT SENTENCE TRANSFORMER MODELS (IN SECONDS).

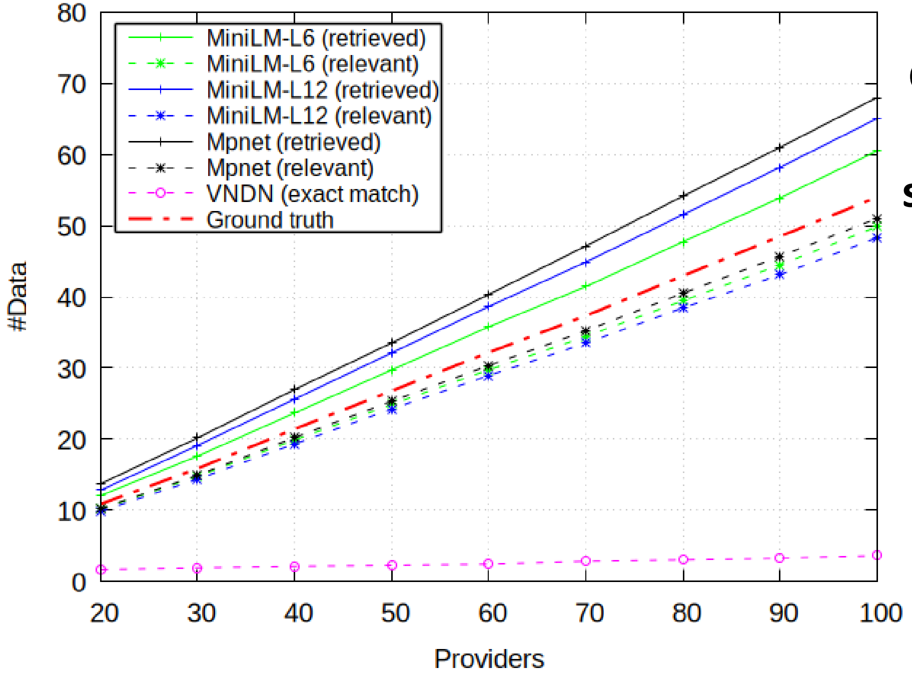
Model	Embedding Time	Matching Time	Total Time
Mpnet	0.0340	0.1601	0.1941
MiniLM-L12	0.0131	0.0419	0.055
MiniLM-L6	0.0095	0.0256	0.0351

Too restrictive, resulting in under-collection

Th=0.9



Th=0.65



Good balance among selectivity and sensitivity

M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm'26

Conclusions

- **NDN for AI**
  - NDN is a natural communication substrate for distributed AI at the edge: its core features (named data, in-network caching, stateful forwarding) directly address networking bottlenecks
  - Early results confirm gains over TCP/IP under realistic lossy wireless conditions for FL and split inference applications
- **AI for NDN**
  - Semantic intelligence embedded in the NDN forwarding fabric overcomes the fundamental naming heterogeneity problem in open, dynamic environments
  - Early results showed flexible data discovery in vehicular crowdsensing without any prior naming agreement

M. Amadeo, C. Campolo, A. Molinaro «NDN for AI and AI for NDN: challenges, opportunities and early results» NDNcomm’26

Future works

- **NDN for AI**
  - Naming schemes for AI artifacts: versioning, model lineage, and parameter provenance
  - Adaptive forwarding strategies for split inference: how should the NDN FIB/PIT be updated dynamically as the splitting strategy changes across nodes?
  - Caching of inference outputs or model updates
- **AI for NDN**
  - Strategies for ground-truth validation in real scenario
  - Multi-hop forwarding with semantic awareness
  - Semantic-based caching

Main references

NDN for AI

- C. Campolo, G. Lia, M. Amadeo, G. Ruggeri, A. Iera, A. Molinaro "Towards named AI networking: Unveiling the potential of NDN for edge AI" *International Conference on Ad-Hoc Networks and Wireless* (pp. 16-22), 2020.
- M. Amadeo, C. Campolo, A. Molinaro, G. Ruggeri, G. Singh "Mitigating the communication straggler effect in federated learning via named data networking" *IEEE Communications Magazine* 62.11 (2024): 92-98.
- M. Amadeo, C. Campolo, A. Molinaro, G. Ruggeri, G. Singh "In-Network Split Inference with Named Data Networking" *IEEE NetSoft workshop* 2025.
- M. Amadeo, C. Campolo, A. Molinaro, G. Ruggeri "In-Network Split Inference with Named Data Networking under Lossy Edge Connectivity" *IEEE 21st International Conference on Network and Service Management (CNSM)*, 2025.
- M. Amadeo, C. Campolo, A. Molinaro, G. Ruggeri "Improving communication performance of Federated Learning: A networking perspective " *Computer Networks* 267 (2025): 111353.

Main references

AI for NDN

- M. Amadeo, S. Serrano, A. Molinaro, M. Nitti, G. Ruggeri "Enhancing IoT Service Discovery through Semantic Name-based Forwarding " *IEEE Internet of Things Magazine,* 2026.
- M. Amadeo, G. Ruggeri, A. Molinaro, M. Nitti, S. Serrano "Service Provisioning in Digital Twin Networks with Semantic-Aware Information Centric Networking" *IEEE CCNC 2026.*
- M. Amadeo, G. Ruggeri, A. Molinaro, S. Serrano "Collective Semantic Intelligence for Vehicular Crowdsensing over Named Data Networking", *accepted at DCOSS-IoT Workshops 2026.*
- M. Amadeo, C. Campolo, S. Serrano, A. Molinaro, G. Ruggeri "DT-Assisted Vehicular Crowdsensing Through Semantic-Aware NDN" *IEEE International Mediterranean Conference on Communications and Networking (MeditCom),* 2025.

Thank you!

Any questions?

marica.amadeo@unime.it