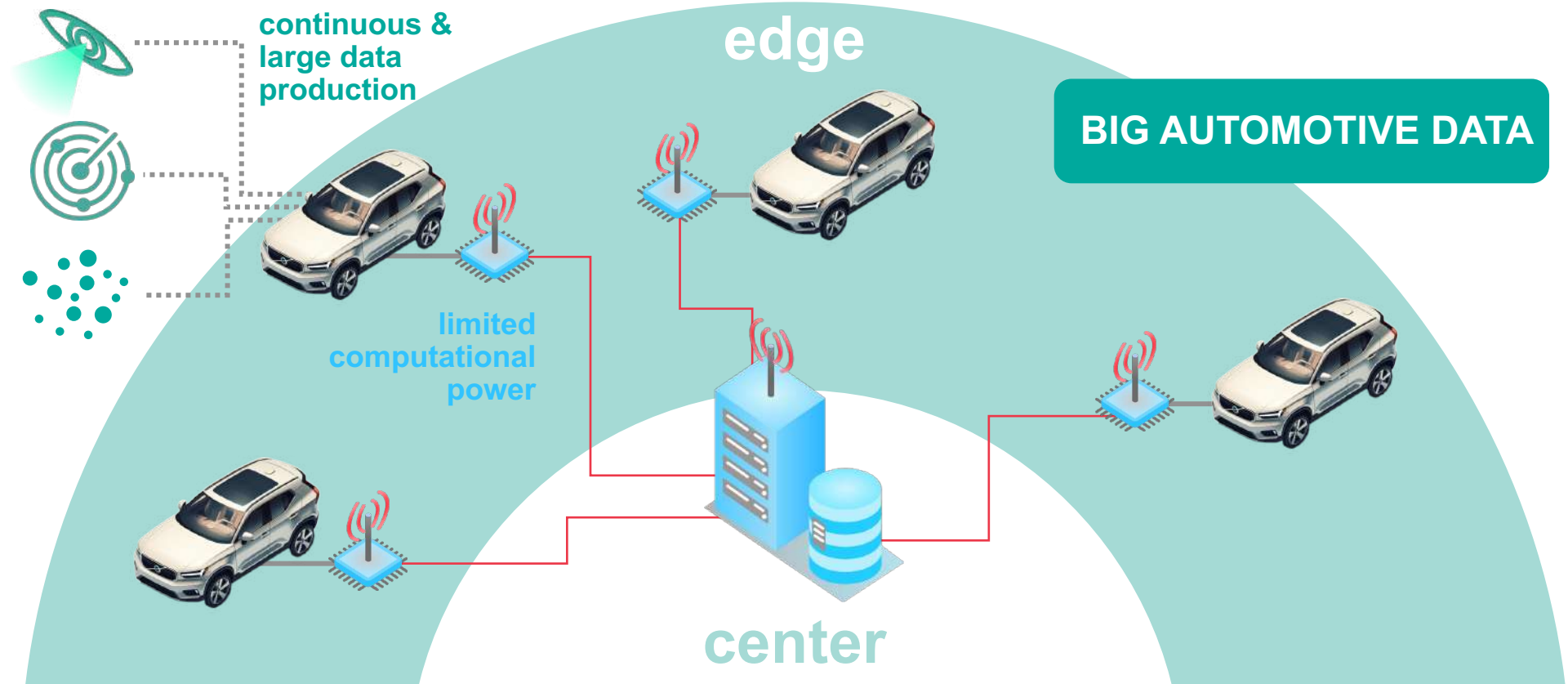
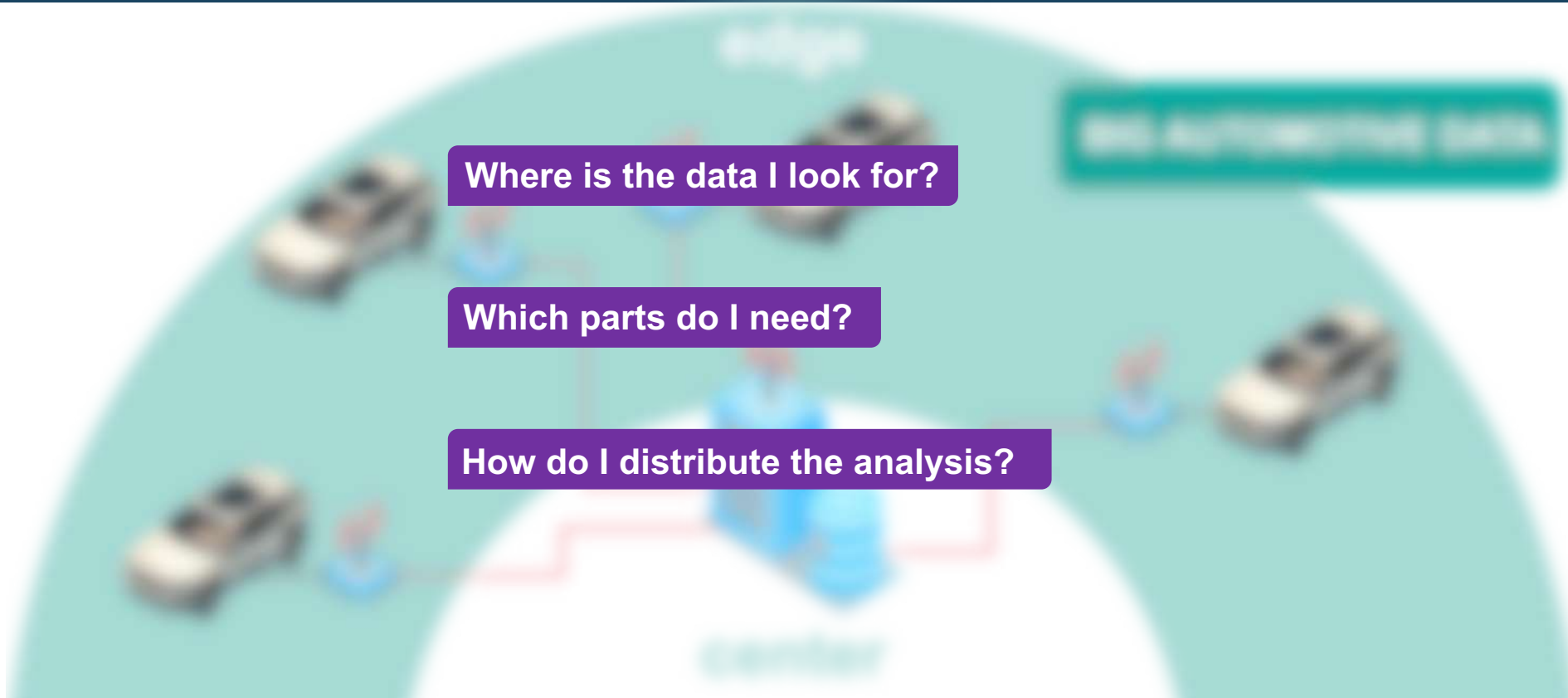
Research Summary: Enhancing Localization, Selection, and Processing of Data in Vehicular Cyber-Physical Systems

Bastian Havers-Zulka (Volvo Car Corporation & Chalmers University)
Vincenzo Gulisano (Chalmers University)
Marina Papatriantafilou (Chalmers University)

Vehicular Cyber-Physical System (VCPS)

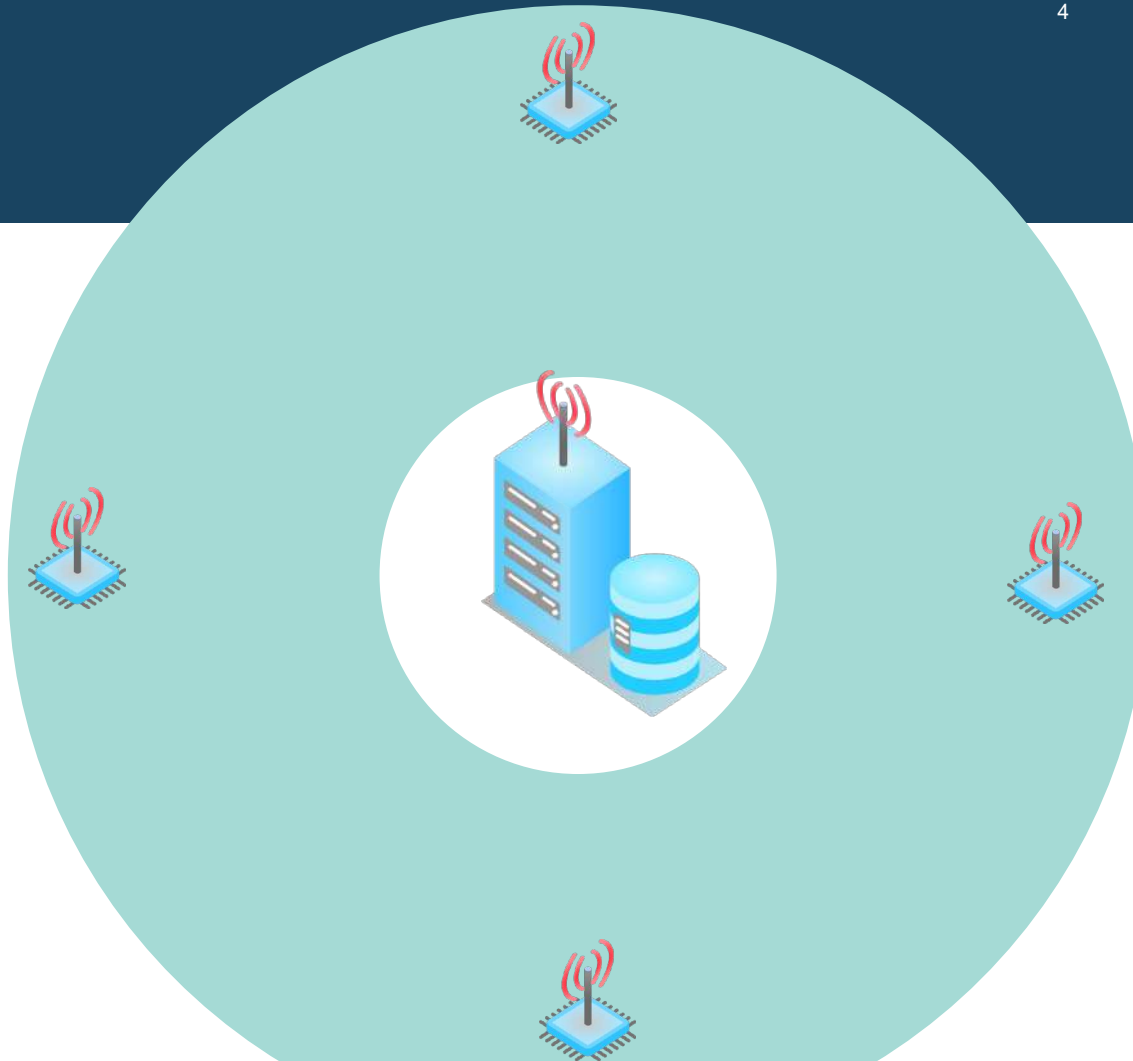


Vehicular Cyber-Physical System (VCPS)



Data Localization

data localization | data gathering

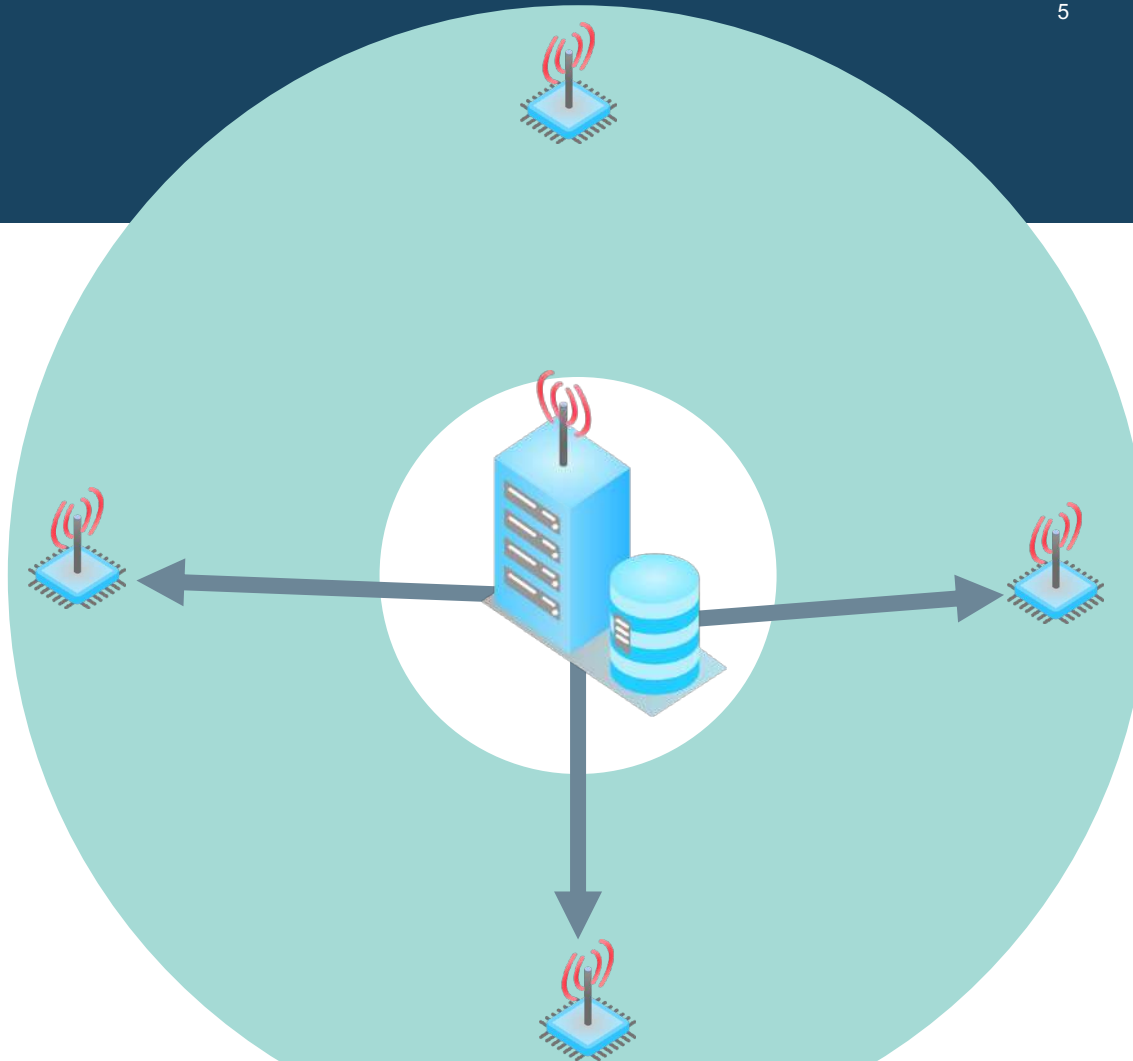


Data Localization

data localization | data gathering

1. **request**

”find **n=2** vehicles that have driven on **parking lots today**”



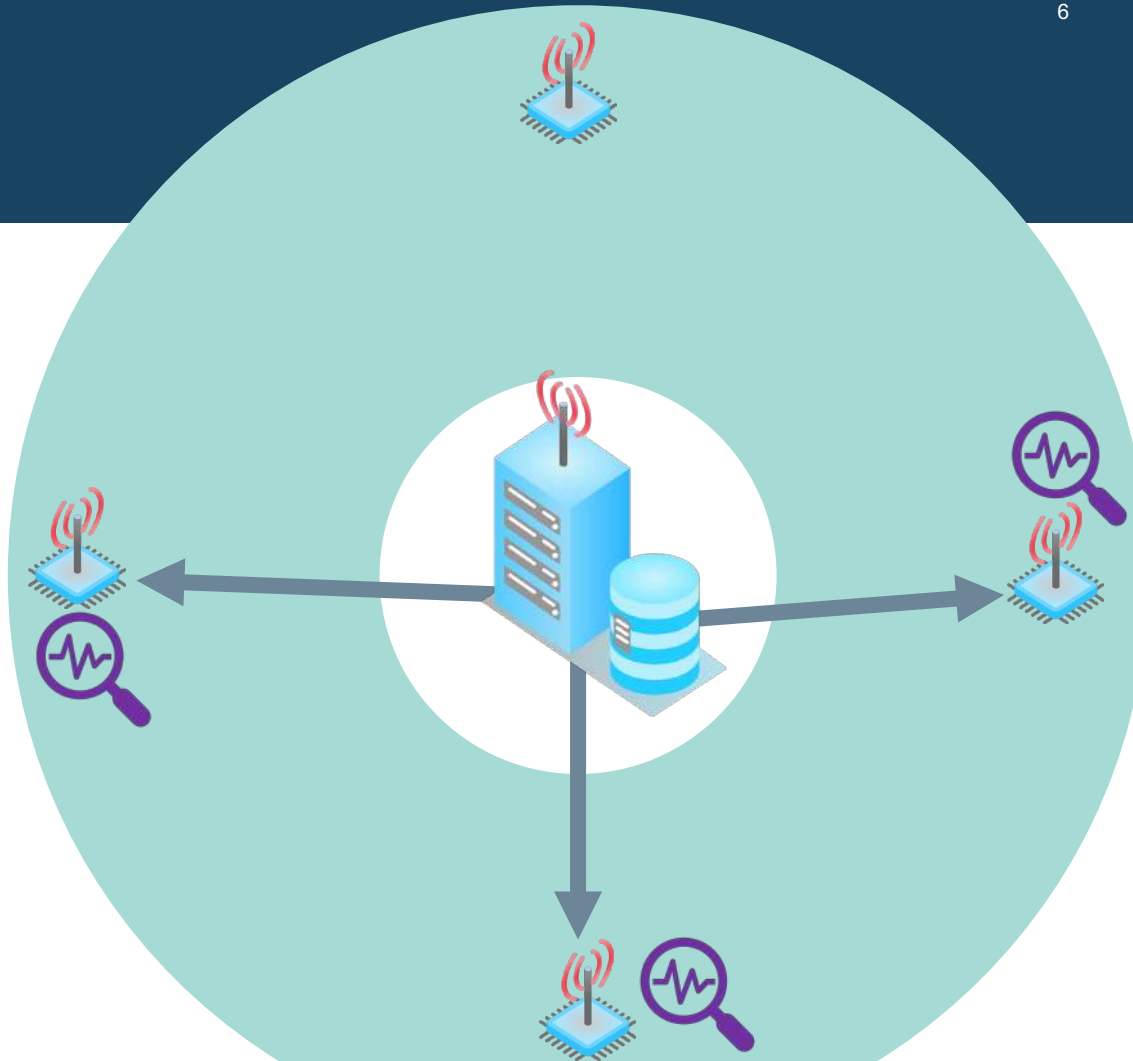
Data Localization

data localization | data gathering

1. **request**

”find **n=2** vehicles that have driven on **parking lots today**”

2. **local** check



Data Localization

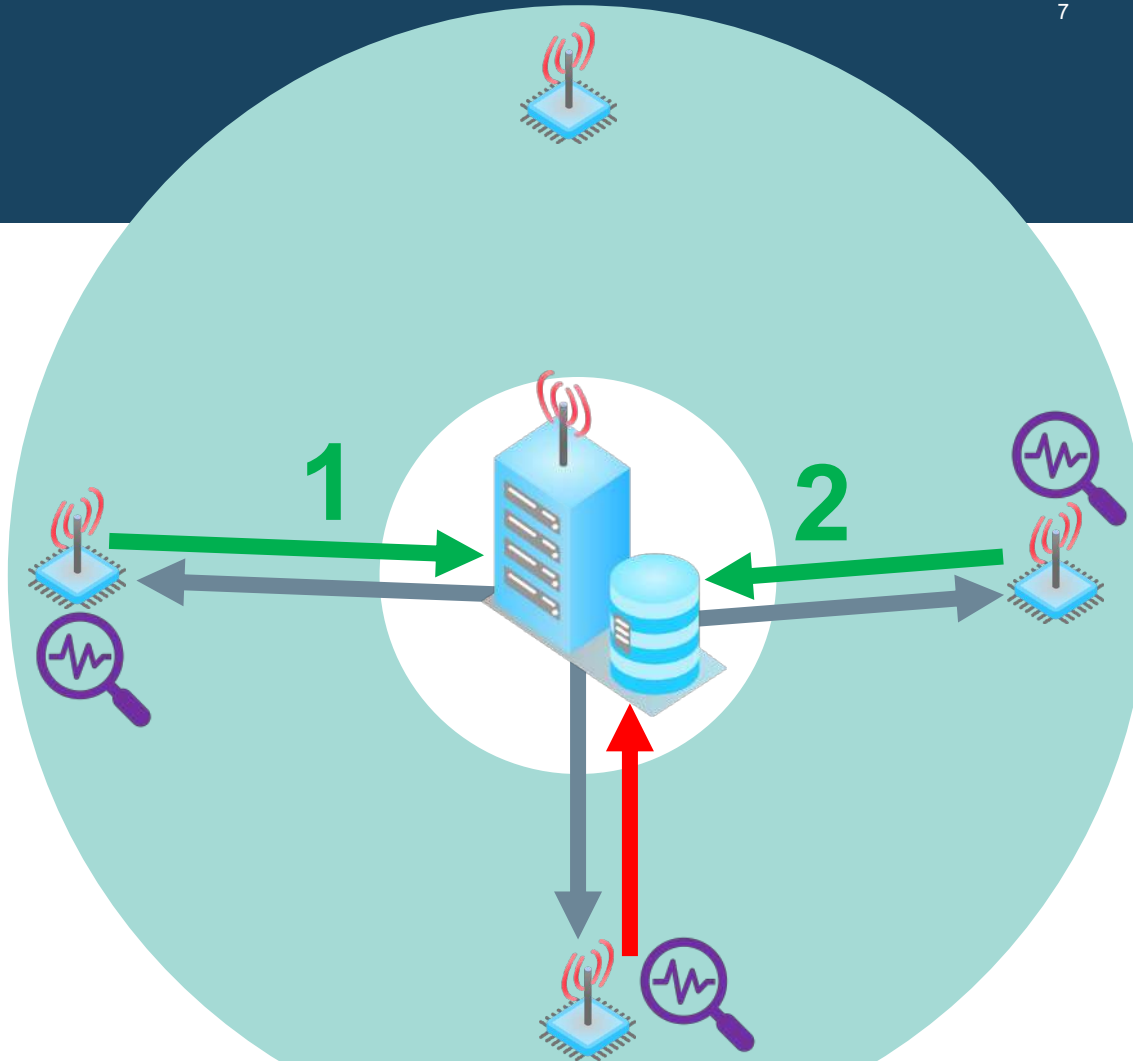
data localization | data gathering

1. **request**

”find **n=2** vehicles that have driven on **parking lots today**”

2. **local** check

3. reply

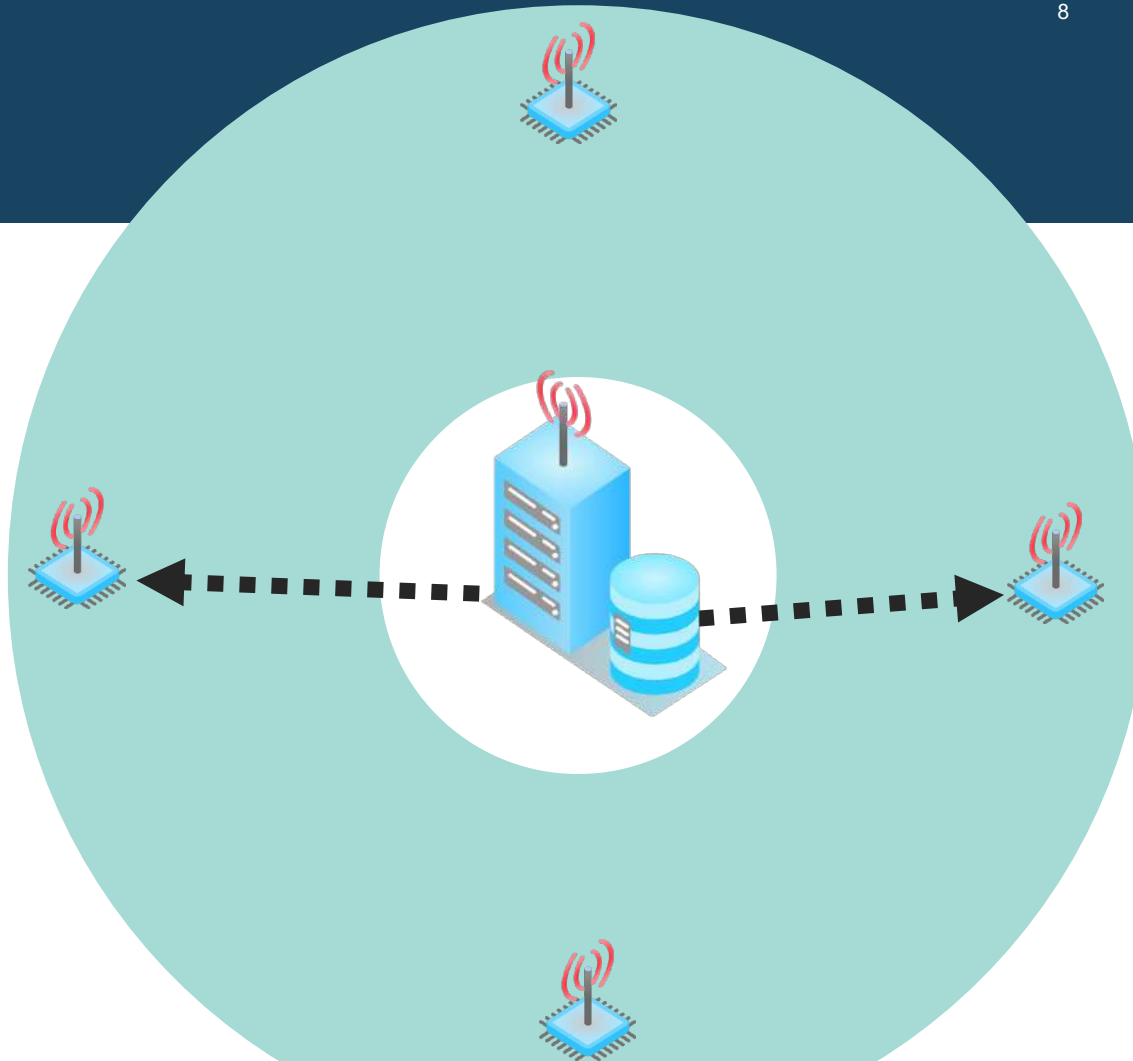


Data Localization

data localization | data gathering

”find **n=2** vehicles that have driven on **parking lots today**”

1. **request**
2. **local** check
3. reply
4. targeted gathering / edge assignment



Contribution

1 Data Distribution

2 Fleet is dynamic

Immediate

ask all, single round

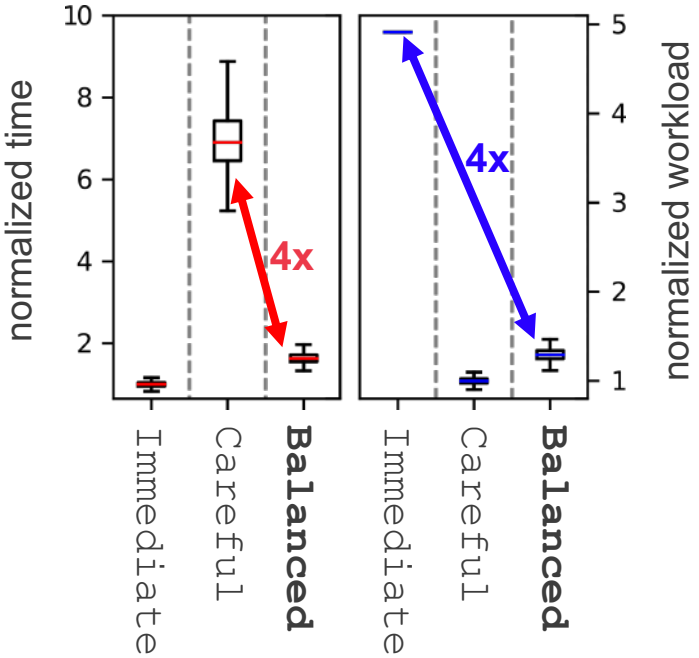
Balanced

- round-based
- **estimate *yes-ratio*** to scale # of requests
- **don't wait** for all answers

Careful

ask 1-by-1

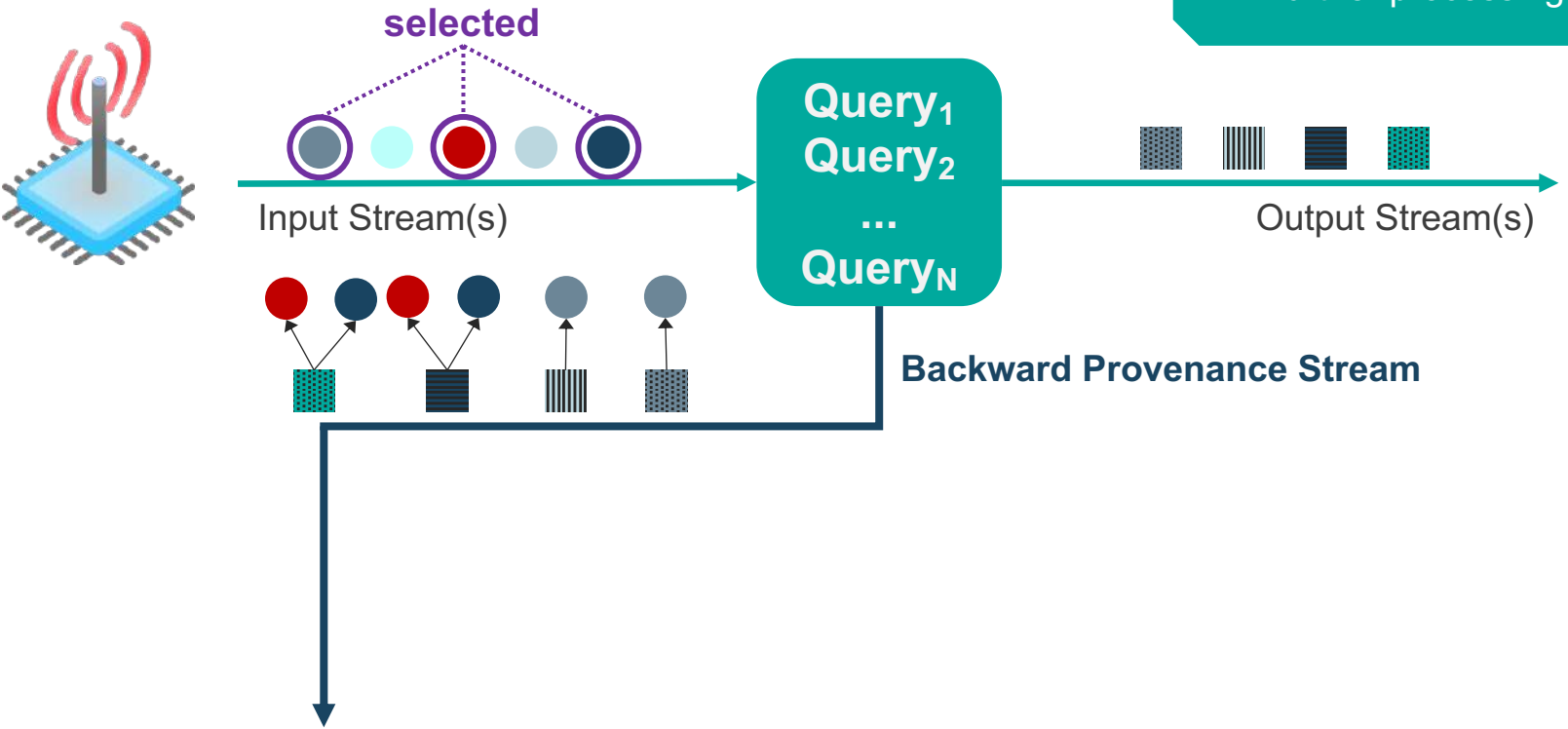
Evaluation Highlights



- **low-powered** hardware
- **two datasets**
- up to **15 queries** simultaneously
- **4x faster, 4x less workload** than baselines

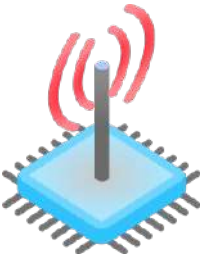
Backward Provenance

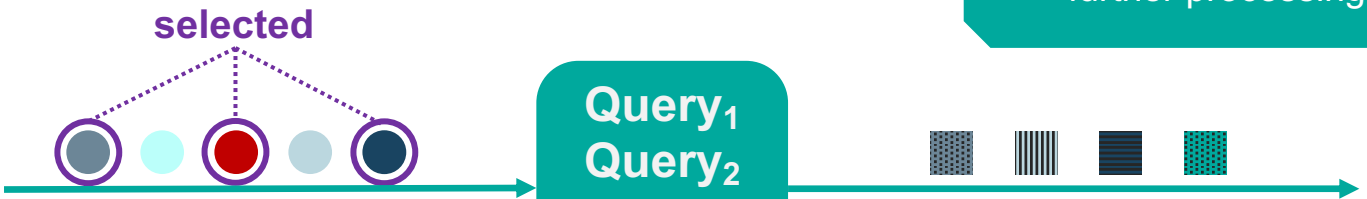
Goals:
1. reduce data duplication
2. indicate once data is ready for further processing

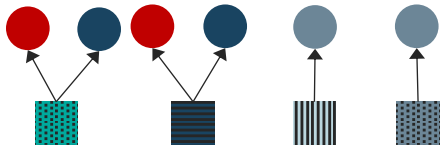


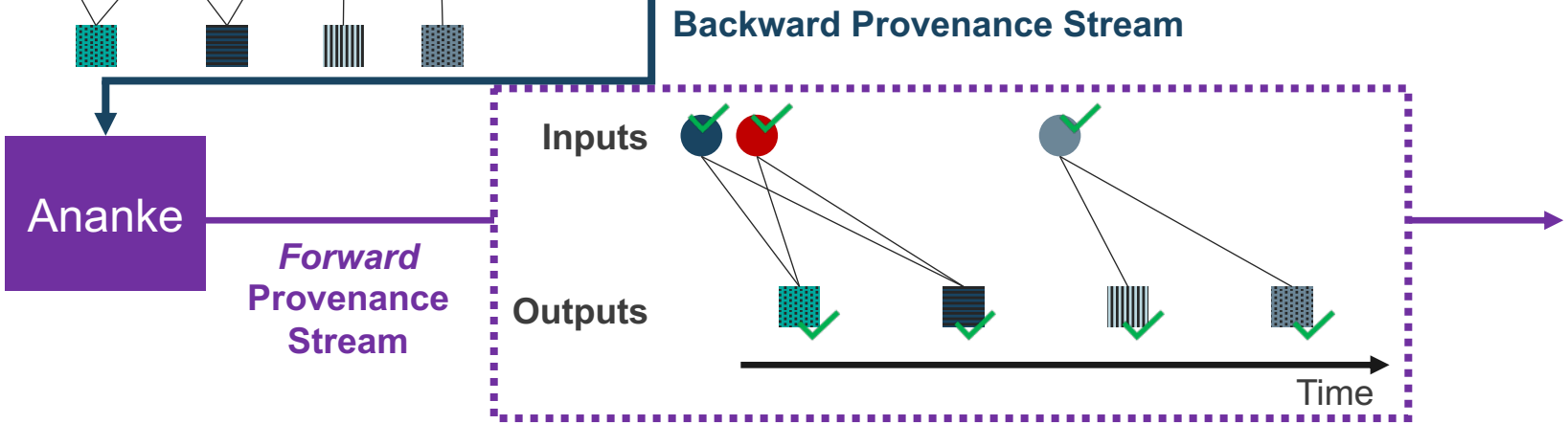
Forward Provenance

Goals:
1. reduce data duplication
2. indicate once data is ready for further processing

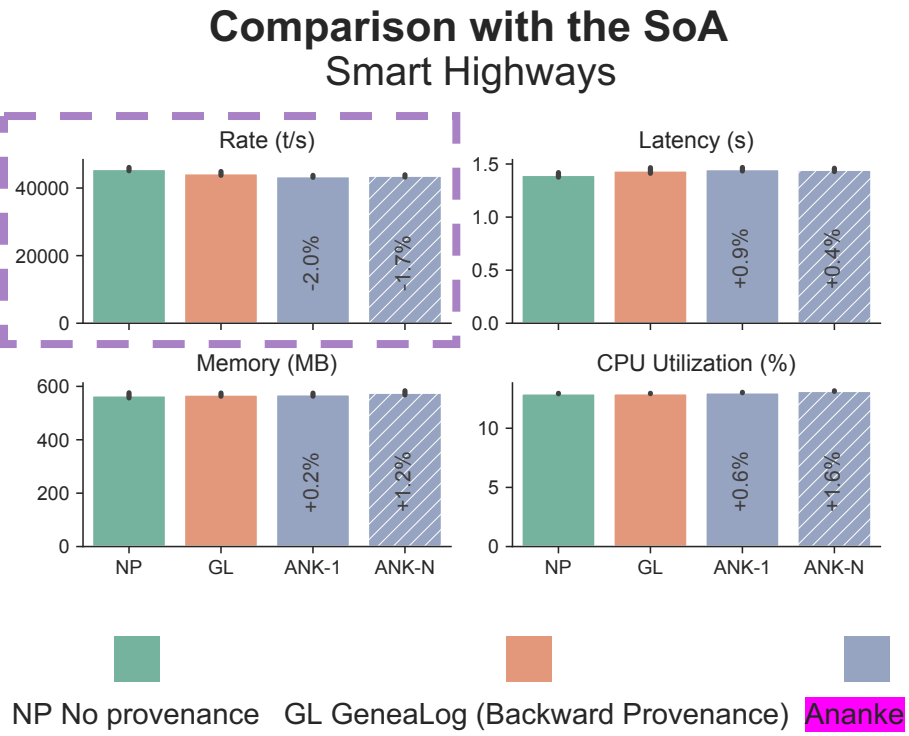








Evaluation Highlights



- **low-powered + high-powered** hardware
- **5 queries**
- **Apache Flink**
- frequently **less than 5% overhead** over SoA
- ANK-N allows to use parallel architectures
- **significantly better performance** than on-demand database techniques

DRIVEN framework

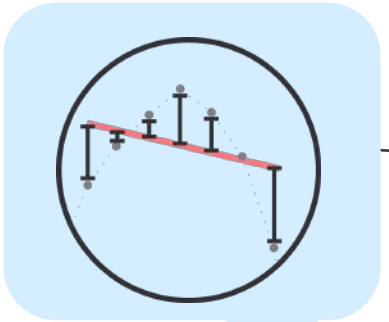
1

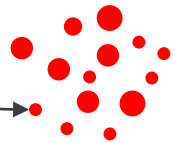
data gathering: slow + costly transmission

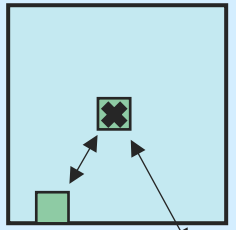
Goal:
Jointly address gathering & analysis challenge while leveraging the edge

2

batch analysis: result latency







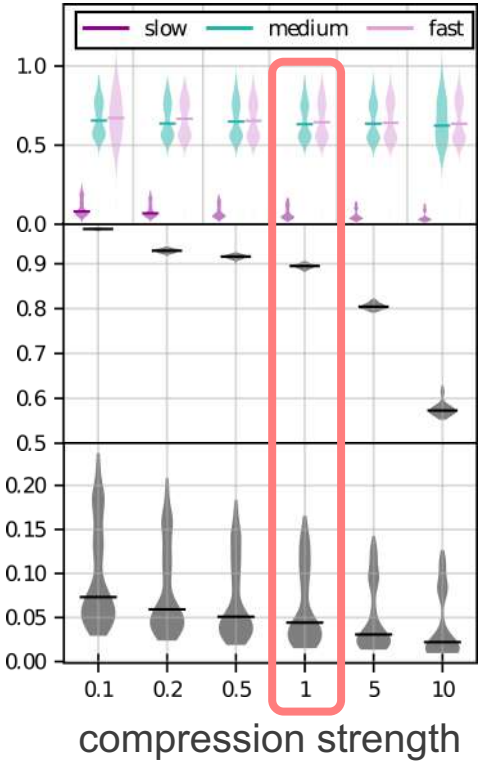
lossy streaming PLA

streaming clustering

$$compression\ ratio = f(error)$$

Najdataei, H., et al. "Continuous and parallel lidar point-cloud clustering." *ICDCS 38*. IEEE, 2018.

Evaluation Highlights



gathering time ratio

clustering accuracy

compression ratio

- **low-powered** hardware
- **heterogeneous** data
- **significant data reductions**
- **tuneable trade-offs**, e.g. **90% accuracy**, **1.6 times faster**, **20 times smaller**

Conclusions

→ edge compression **jointly leverages** Stream and edge processing;

→ can deliver 90% data savings on LiDAR at small accuracy loss

→ **Forward Provenance** can efficiently aid in **Data Selection** to further reduce data amounts

Further results from our work:

→ **Data Localization** can be performed with **near-baseline performance**

→ Novel communication schemes can reduce data volumes in **Distributed Machine Learning** on the edge