UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

# A Location-Allocation model for Fog Computing Infrastructures


Thiago Alves de Queiroz
*IMT, Federal University of Goiás*

Claudia Canali,
Riccardo Lancellotti

*DIEF, University of Modena and Reggio Emilia*

Manuel Iori

*DISMI, University of Modena and Reggio Emilia*

# New challenges

- New paradigm: Smart cities large scale sensing applications
- Several fields of application:
  - Urban applications
  - Industrial
  - Automotive
  - Healthcare
  - ...
- New scenarios: Cyber-physical systems
  - Geographically distributed sensors
  - Huge amount of information produced

# New challenges

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

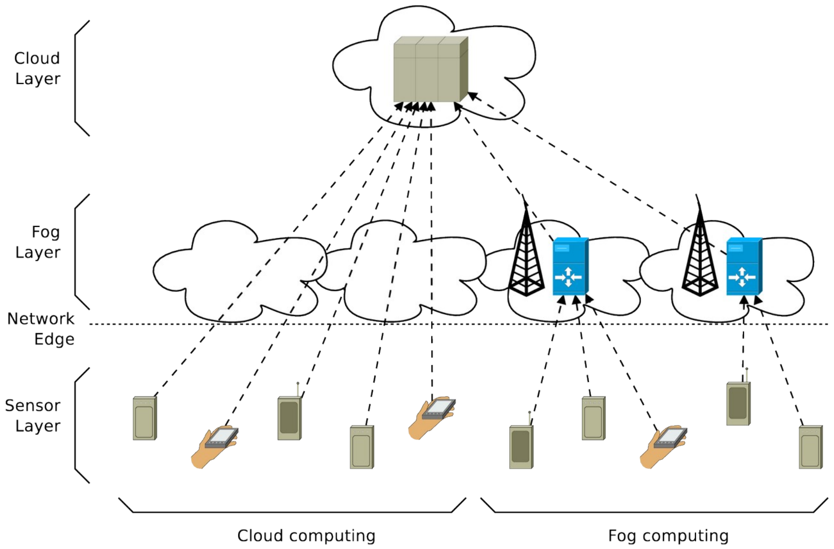
→ New requirements for the infrastructure

- Scalability challenge
  - Huge amount of data to transfer and process
  - Geographically distributed systems
  - Example: CPU- and bandwidth-bound applications
- Low latency challenge
  - Support for real time applications
  - Example: latency-bound applications
- Cloud computing is not enough
- *(5G alone is not an answer)*

# Pros and Cons of Fog

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

- Benefits of Fog computing
- Scalability:
  - Pre-processing offloaded to fog nodes
  - Less strain on Cloud network links
- Latency:
  - Latency-critical tasks offloaded to Fog
  - Fog nodes are closer to the edge




- New open issues:
  → new Fog infrastructure
  - Fog node deployment
  - Sensors-to fog mapping
- Joint problem

# Our contribution

- Model for the design of Fog infrastructures
  - Based on location-allocation optimization problem
- Model decisions:
  - How many fog nodes do we need?
  - Which Fog nodes (among a set) turn on?
  - How to map sensors over fog nodes?
- Double optimization goal
  - Reduce infrastructure cost
  - Optimize performance
- Use of SLA constraints

# Notation

UNIMORE UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

| | |
|---|---|
| **Model parameters** | |
| $\mathcal{S}$ | Set of sensors |
| $\mathcal{F}$ | Set of fog nodes |
| $\mathcal{C}$ | Set of cloud data centers |
| $\lambda_i$ | Outgoing data rate from sensor $i$ |
| $\lambda_j$ | Incoming data rate at fog node $j$ |
| $1/\mu_j$ | Processing time at fog node $j$ |
| $\delta_{ij}$ | Communication latency between sensor $i$ and fog $j$ |
| $\delta_{jk}$ | Communication latency between fog $j$ and cloud $k$ |
| $c_j$ | Cost for locating a fog node at position $j$ (or for keeping the fog node turned on) |
| **Model indices** | |
| $i$ | Index for a sensor |
| $j$ | Index for a fog node |
| $k$ | Index for a cloud data center |
| **Decision variables** | |
| $E_j$ | Location of fog node $j$ |
| $x_{ij}$ | Allocation of sensor $i$ to fog $j$ |
| $y_{jk}$ | Allocation of fog node $j$ to cloud $k$ |

# Optimization problem

UNIMORE UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

- Objective function
  - → Cost for fog nodes
  - → Response time
- Contributions to response time:
  - Sensor → Fog avg net delay
  - Fog → Cloud avg net delay
  - Fog processing time
- Caveat: definition of $\lambda_j$
- Main constraints:
  - Response time < SLA
  - Load on nodes

$$C = \sum_{j \in \mathcal{F}} c_j E_j$$

$$T_R = T_{netSF} + T_{netFC} + T_{proc}$$

$$T_{netsf} = \frac{1}{\sum_{i \in \mathcal{S}} \lambda_i} \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{F}} \lambda_i x_{i,j} \delta_{i,j}$$

$$T_{netfc} = \frac{1}{\sum_{j \in \mathcal{F}} \lambda_j} \sum_{j \in \mathcal{F}} \sum_{k \in \mathcal{C}} \lambda_j y_{j,k} \delta_{j,k}$$

$$T_{proc} = \frac{1}{\sum_{j \in \mathcal{F}} \lambda_j} \sum_{j \in \mathcal{F}} \frac{\lambda_j}{\mu_j - \lambda_j}$$

$$\lambda_j = \sum_{i \in \mathcal{S}} x_{i,j} \cdot \lambda_i$$

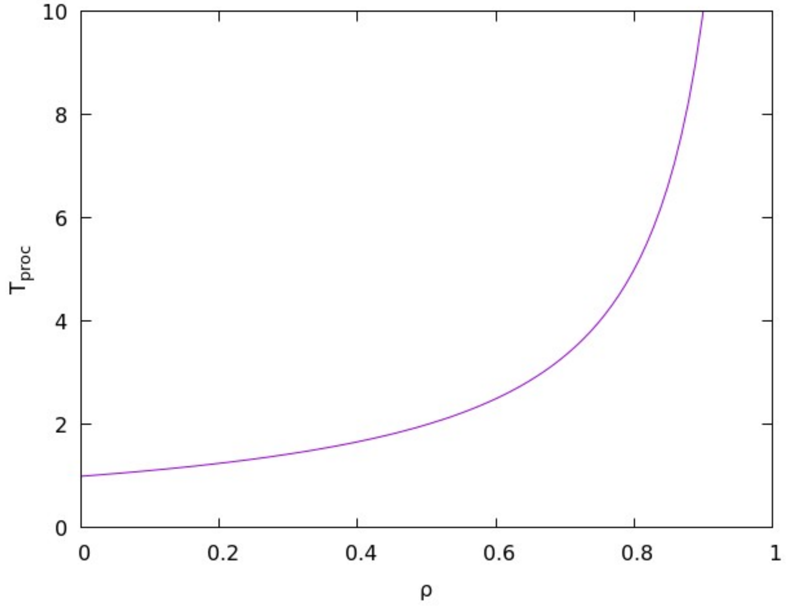
$$T_R \leq T_{SLA}$$

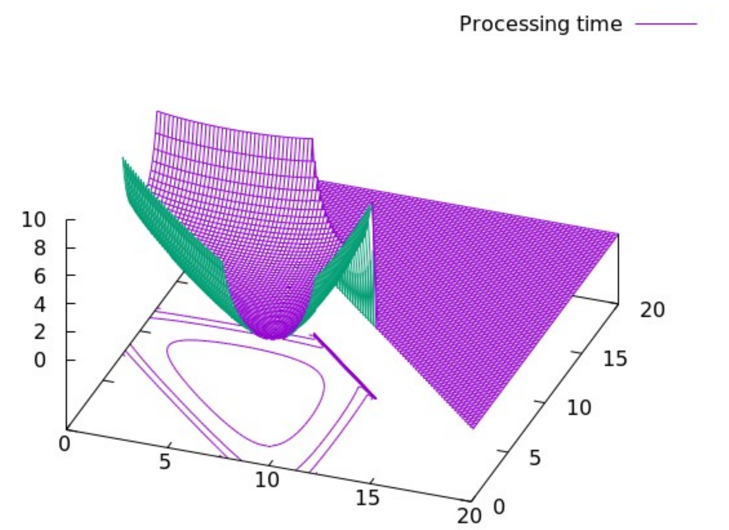
$$\lambda_j < E_j \mu_j, \quad \forall j \in \mathcal{F}$$

# Processing time

- Based on queuing theory
  - M/G/1 models
  - Consistent with PASTA theorem
- Non linear model
- Response time as a function of system load

$$T_{proc} = \frac{1}{\mu - \lambda} = \frac{1}{\mu} \cdot \frac{1}{1-\rho}$$



# Scenario definition

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

- Parameters to describe scenarios
- Average network delay $\delta$
  - Typically set to ~10ms
- Network delay / Processing time balance $\delta\mu$
  - Scenario CPU bound or Network bound
- System load $\rho$
  - Average load of fog nodes

$$\delta = \frac{\sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{F}} \delta_{i,j} + \sum_{j \in \mathcal{F}} \sum_{k \in \mathcal{C}} \delta_{j,k}}{|\mathcal{S}| \cdot |\mathcal{F}| + |\mathcal{F}| \cdot |\mathcal{C}|}$$

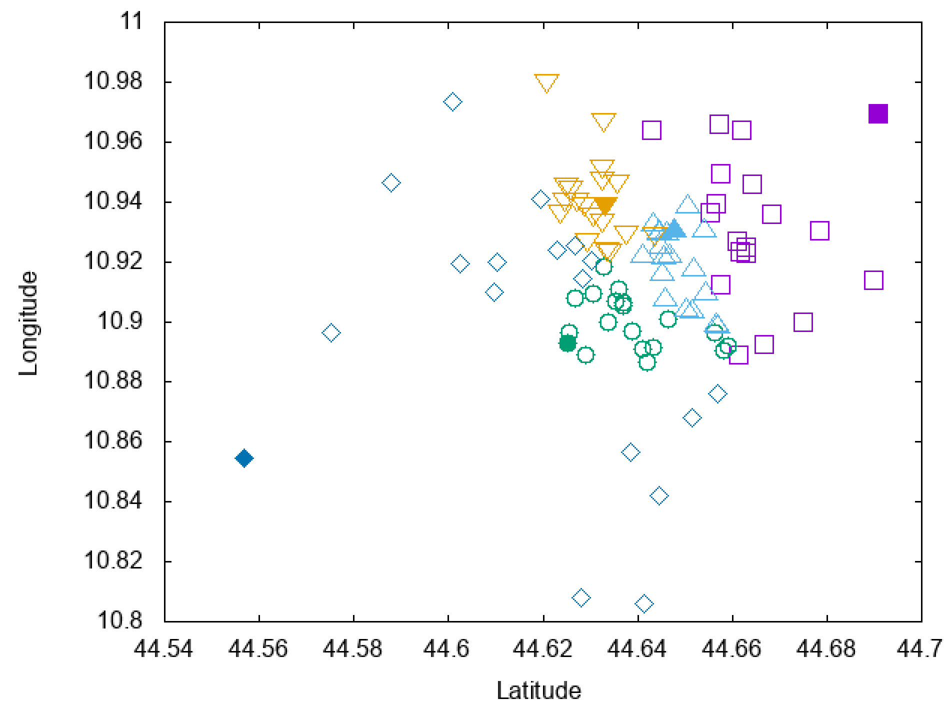
$$\delta\mu = \delta \cdot \frac{\sum_{j \in \mathcal{F}} \mu_j}{|\mathcal{F}|}$$

$$\rho = \frac{\sum_{i \in \mathcal{S}} \lambda_i}{\sum_{j \in \mathcal{F}} \mu_j}$$

# Experimental scenario

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

- Smart City scenario based on real example
  - Italian city (Modena),
  - ~180,000 inhabitants
- Traffic monitoring case
  - Sensors on streets
  - Fog nodes in public buildings
  - LoRa connections
- Evaluation using solver
- Comparison with:
  - Continuous model (no bool)
  - Simplified model ($E_i = 1$)




LocalSolver

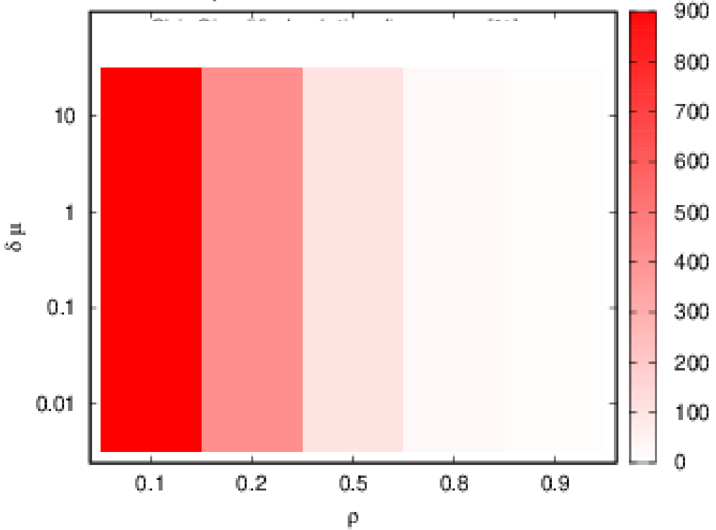
(Ideal lower bound, used as baseline comparison)

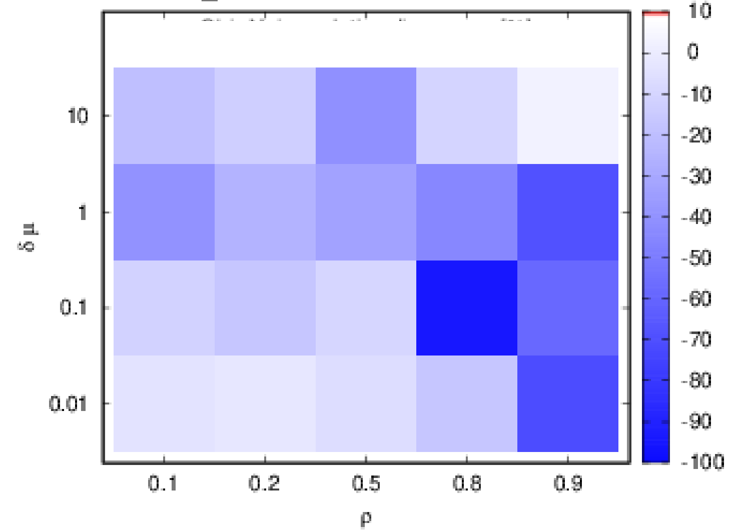
# Experimental results

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

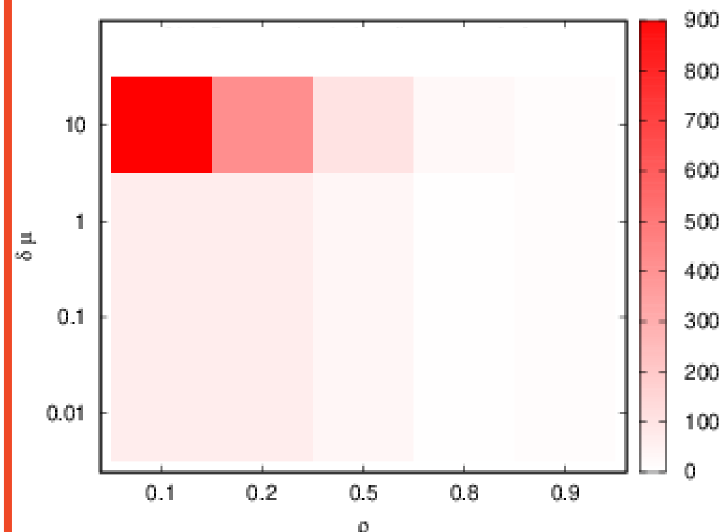
Obj$_1$ divergence [%] | Obj$_2$ divergence [%]

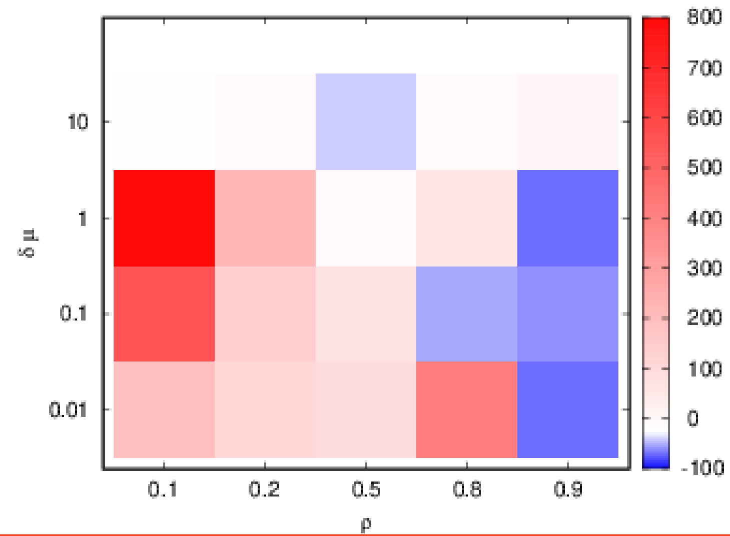
Simplified Model




Proposed Model



# Conclusions

- Challenges of Fog computing
  - Selection of fog nodes and mapping of sensors
- Contribution: proposal of a model
  - Based on location-allocation optimization problem
  - Dual objective function
  - Non linear problem
- Validation of the model
  - Focus on a realistic scenario
  - Wide range of parameters considered
- Open issues
  - Heuristics (GA, Variable Neighborhood Search)
  - Dynamic scenarios