

API-driven Task Scheduling and Offloading with PULCEO*: An Extension

*Platform for Universal and Lightweight Cloud-Edge Orchestration

Sebastian Böhm · Guido Wirtz

- Distributed Systems Group -

Faculty of Information Systems and Applied Computer Sciences, University of Bamberg



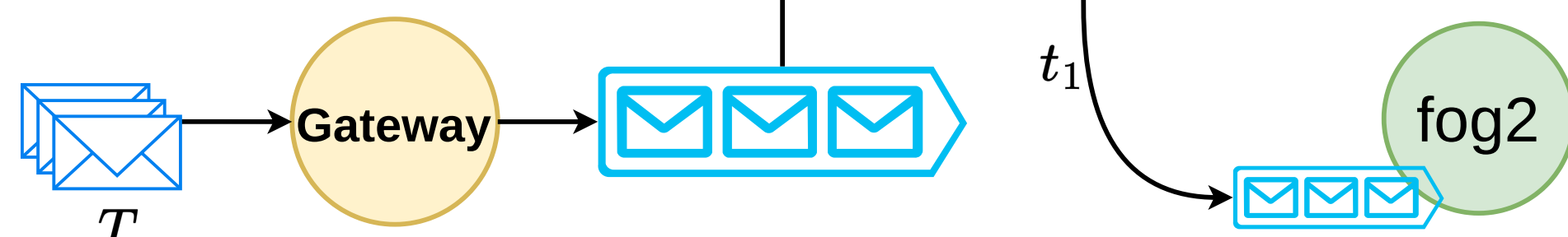
Problem Domain

Task Scheduling / Offloading: Assigning a set of independent tasks T to a set of nodes N , based on properties P and requirements R

$$T = \{P, R\}, N = \{P, R\}$$

$$P(t_i) \subseteq P(n_j)$$

$$R(t_i) \leq R(n_j), \forall k, r_k \in R$$



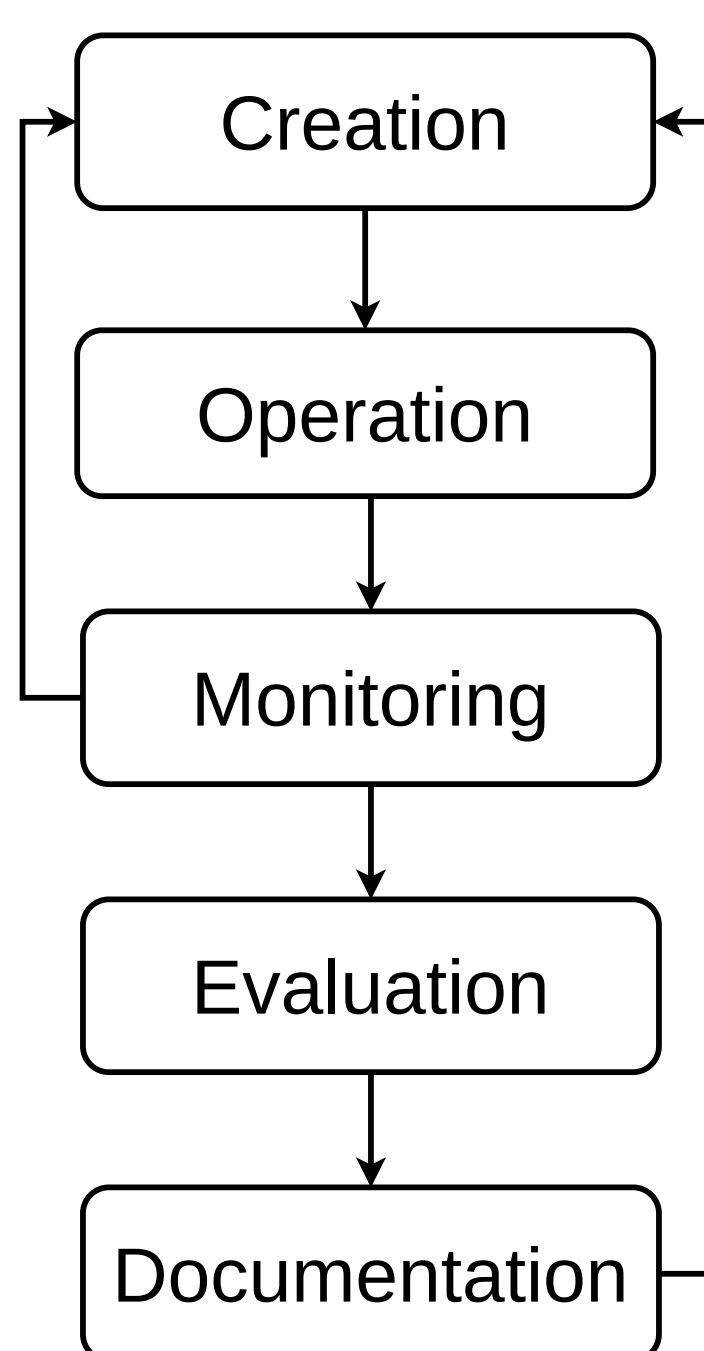
Motivation

Plenty of solutions available

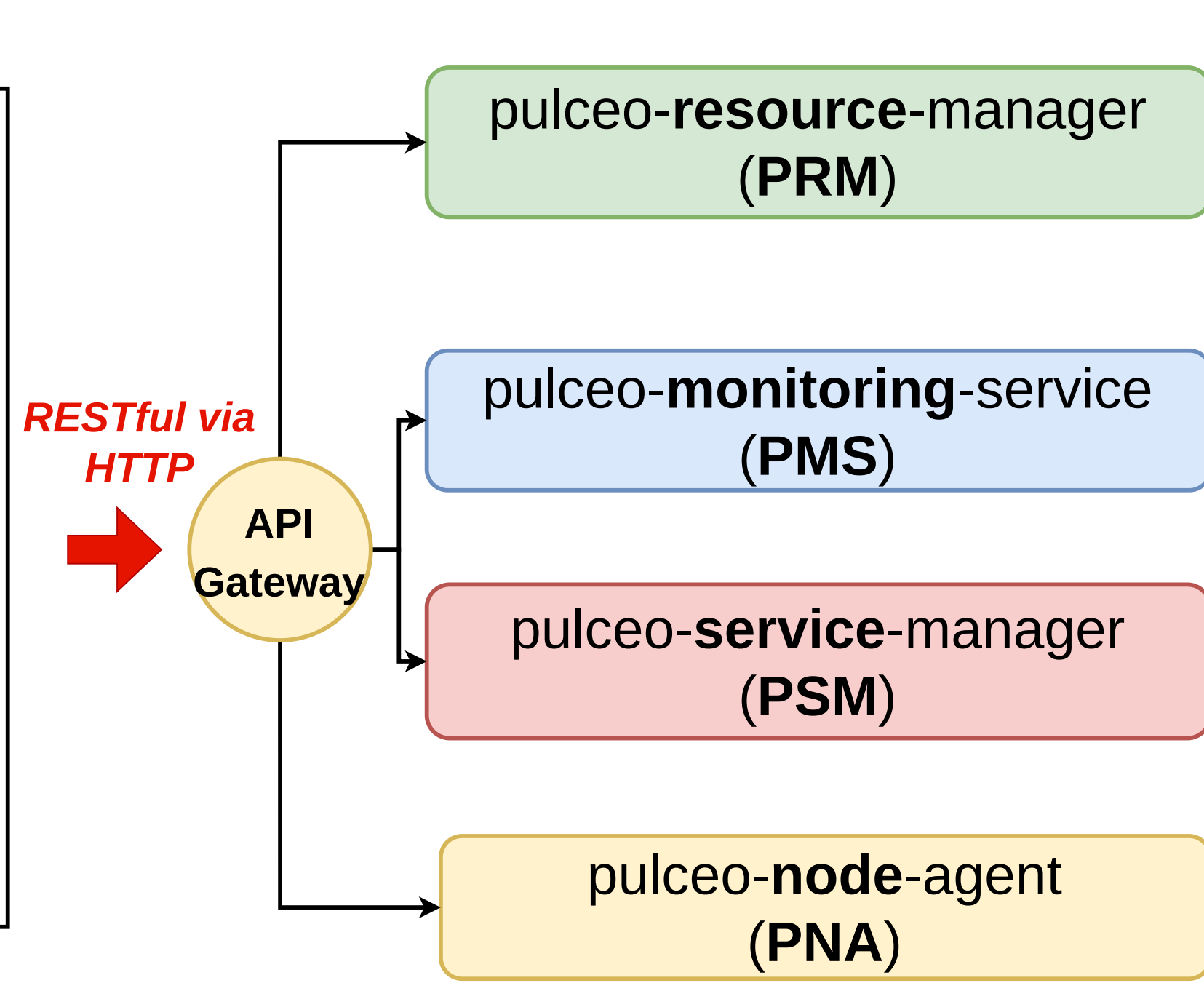
But, mainly evaluated by simulations with no real systems

However, empirical measurements required, like CPU and memory utilization (e.g., fine-tuning for schedulers)

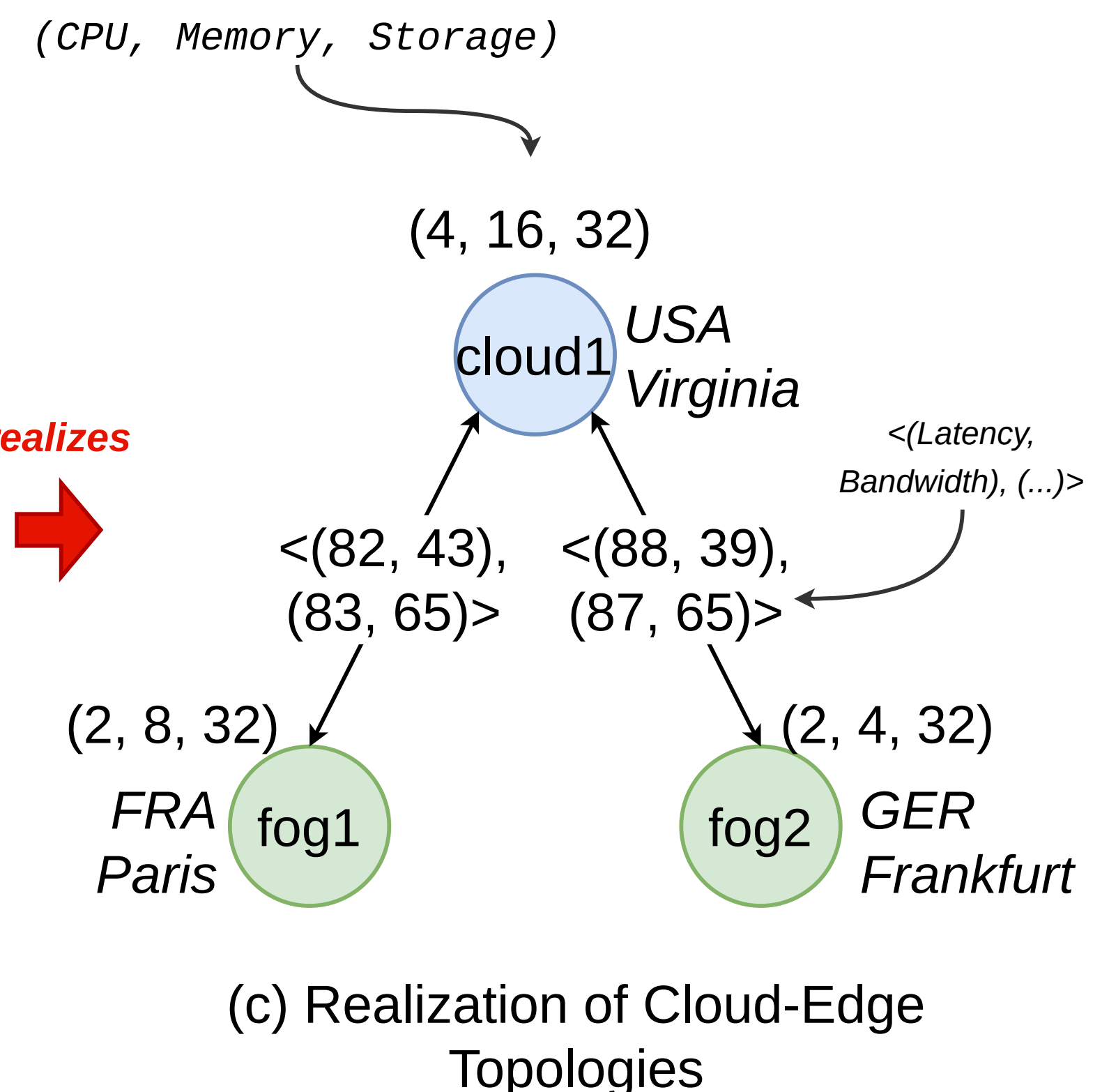
Holistic Cloud-Edge Orchestration with PULCEO



(a) Holistic Management



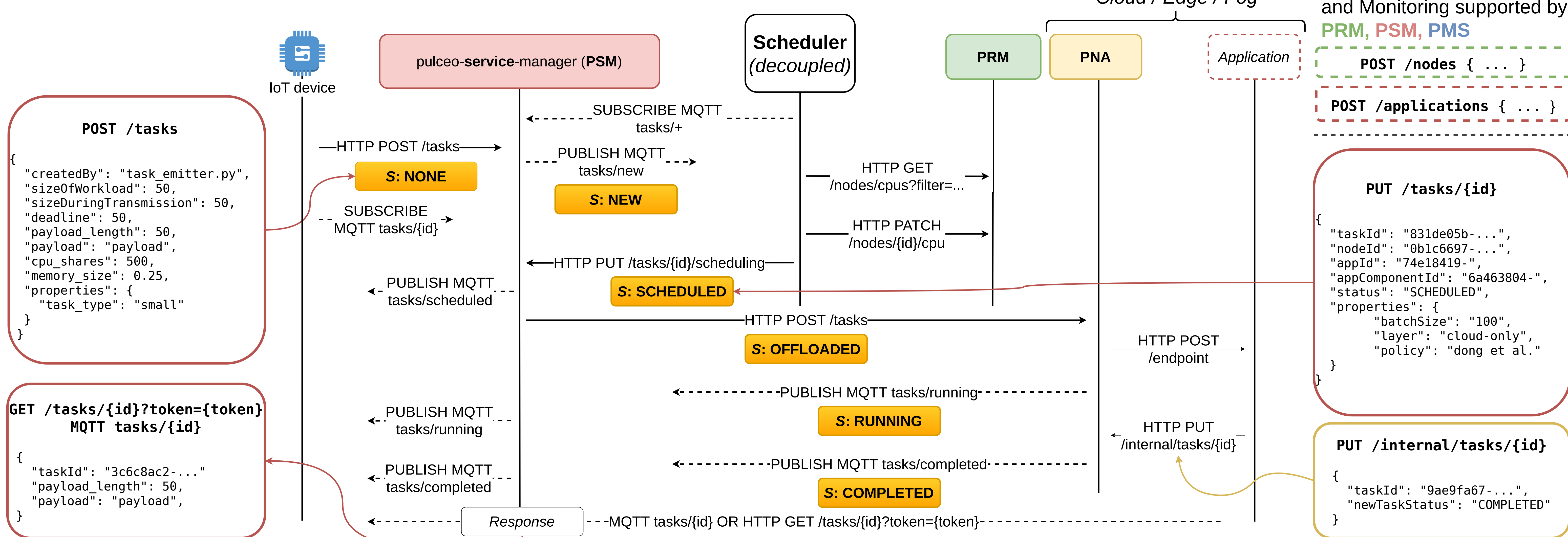
(b) Microservice-based Implementation



(c) Realization of Cloud-Edge Topologies

Solution: Decoupled API-driven Task Scheduling and Offloading in line with Holistic Cloud-Edge Orchestration

② Operation (Task Scheduling and Offloading) supported by PSM via a RESTful HTTP API and MQTT



④ Evaluation with cross-cutting support from all microservices

Task Metrics (related to the internal task processing of PULCEO)

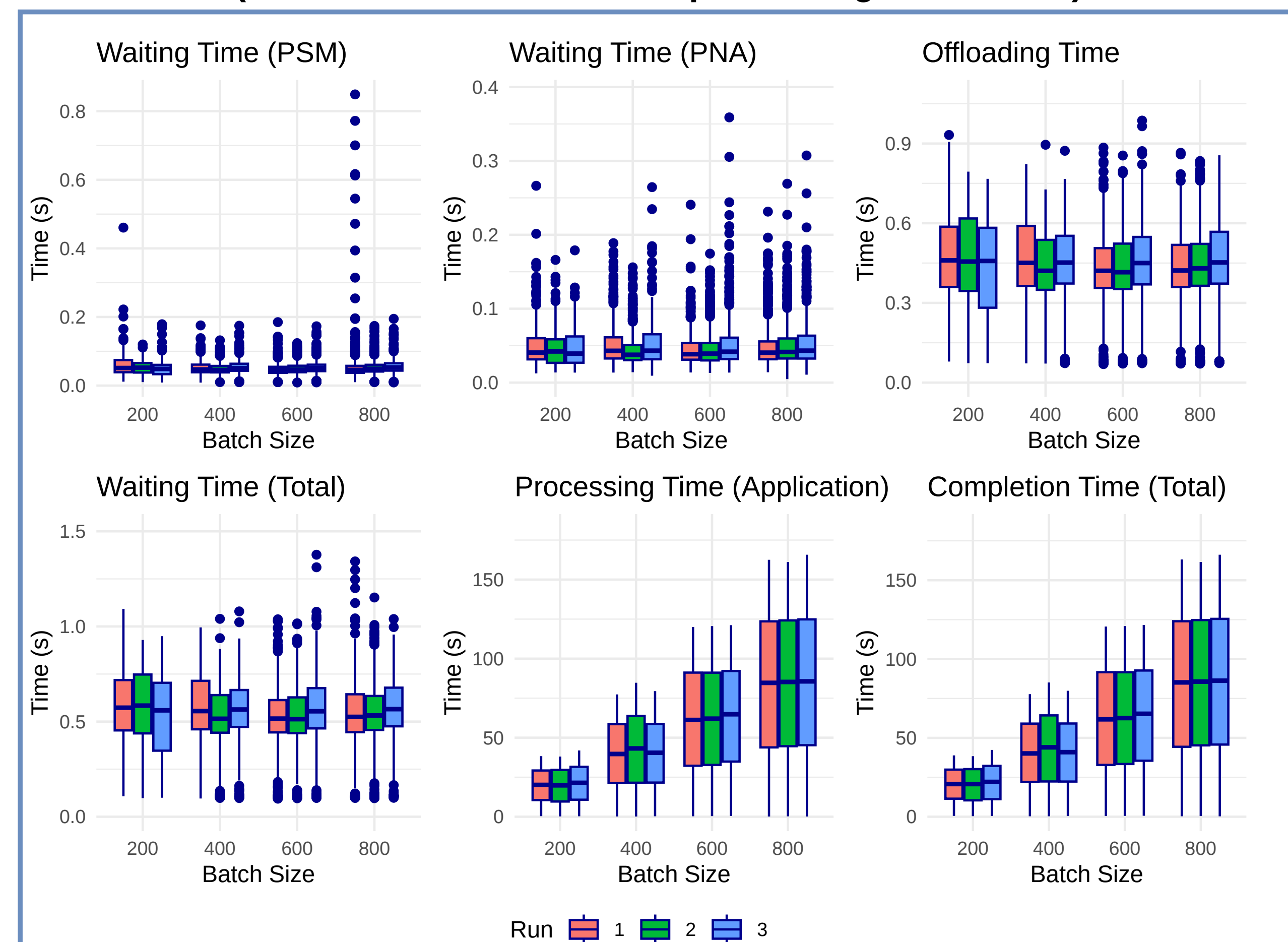
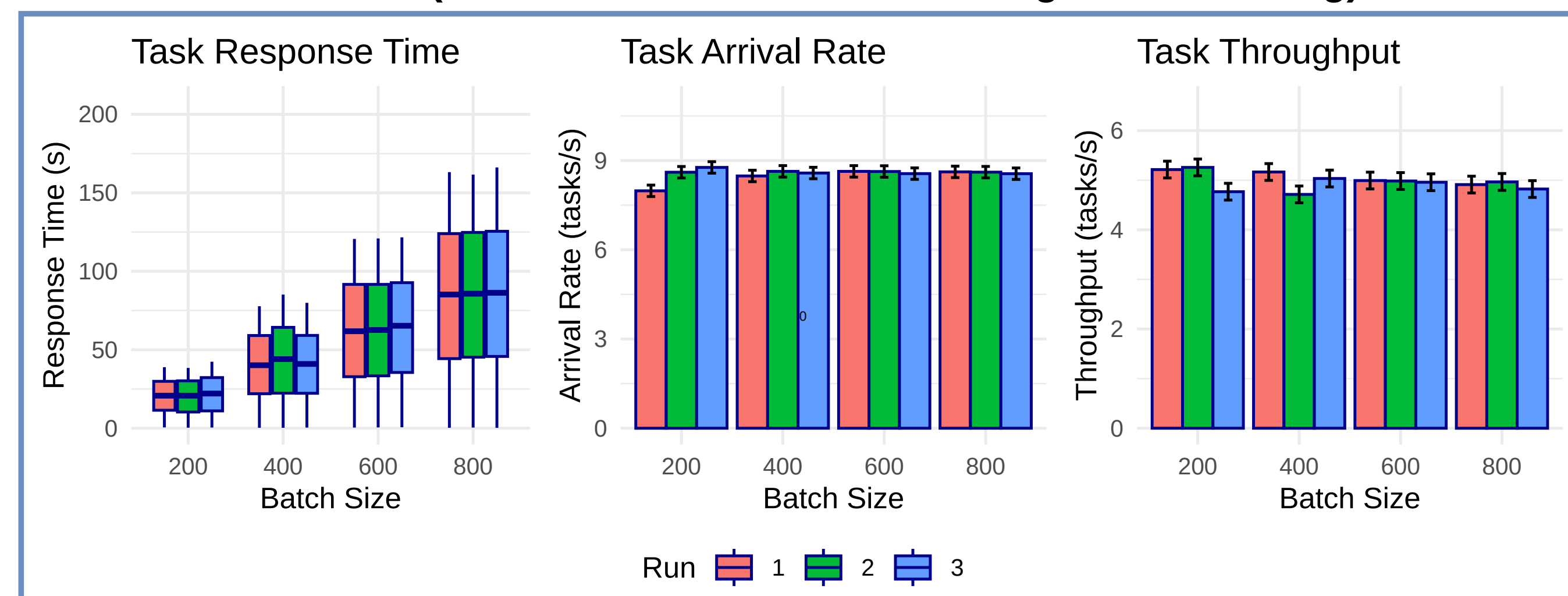


Table 2. Task metric statistics for different batch sizes in seconds (s)

Task metric	S_1	S_2	Batch size			
			200 (μ/σ)	400 (μ/σ)	600 (μ/σ)	800 (μ/σ)
Waiting Time (PSM)	NEW	SCHEDULED	0.05/0.03	0.05/0.02	0.05/0.02	0.05/0.04
Offloading Time	SCHEDULED	OFFLOADED	0.44/0.21	0.45/0.16	0.44/0.16	0.45/0.15
Waiting Time (PNA)	OFFLOADED	RUNNING	0.05/0.03	0.05/0.03	0.05/0.03	0.05/0.03
Waiting Time (Total)	NEW	RUNNING	0.54/0.25	0.55/0.19	0.53/0.18	0.55/0.17
Processing Time (Application)	RUNNING	COMPLETED	20.25/11.33	40.83/22.81	62.09/34.44	84.08/46.55
Completion Time (Total)	NEW	COMPLETED	20.79/11.44	41.38/22.85	62.62/34.47	84.63/46.58

Performance Metrics (related to the overall scheduling and offloading)



⑤ Documentation



- Comprehensibility
- Reproducibility
- Evaluation
- Reporting

Contributions & Limitations

Task Scheduling and Offloading

- Integration of task scheduling and offloading into the holistic cloud-edge orchestration life cycle
- General, universal, and extendable model of tasks
- Seamless technical integration of already available task-processing applications
- Automated standard evaluation with selected metrics for task scheduling and offloading
- Integration of task-processing applications only via HTTP and MQTT

General

- RESTful HTTP API for universal and decoupled cloud-edge orchestration
- Holistic orchestration with creation, operation, monitoring, evaluation, and documentation
- Centralized, cloud-based, and only partly decentralized orchestration

