Adaptive Data Replication in Real-Time Reliable Edge Computing for Internet of Things

[+]Chao Wang, [*]Christopher Gill, [*]Chenyang Lu

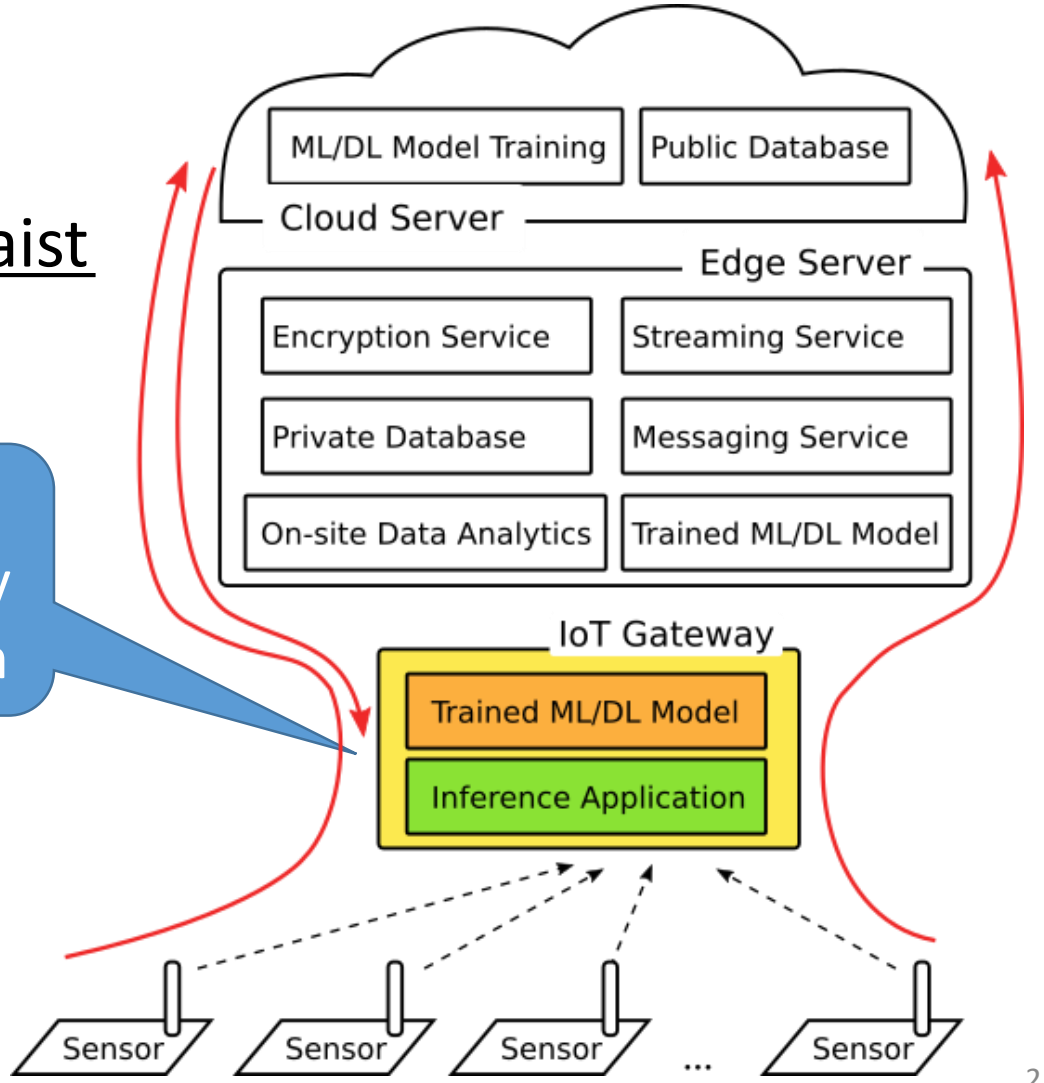
[+] Department of Computer Science and Information Engineering, National Taiwan Normal University

[*] Department of Computer Science and Engineering, Washington University in St. Louis, USA

Challenges in performant IoT edge computing

- Reliable and timely computing
  at a resource-constraint network waist
  an IoT gateway

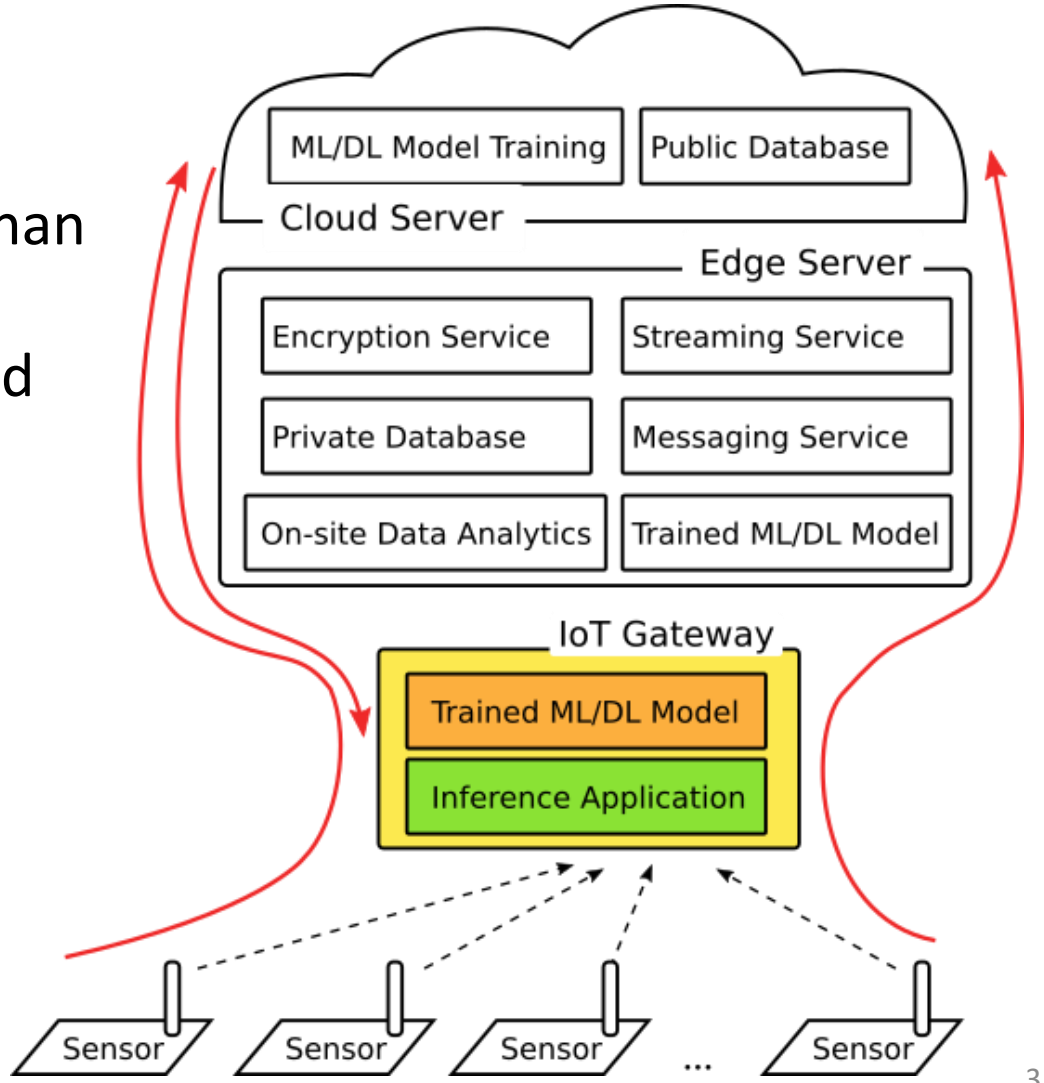
- Single point of failure
- Traffic congestion and delay
- Limited network bandwidth



Example: structural health inference and control

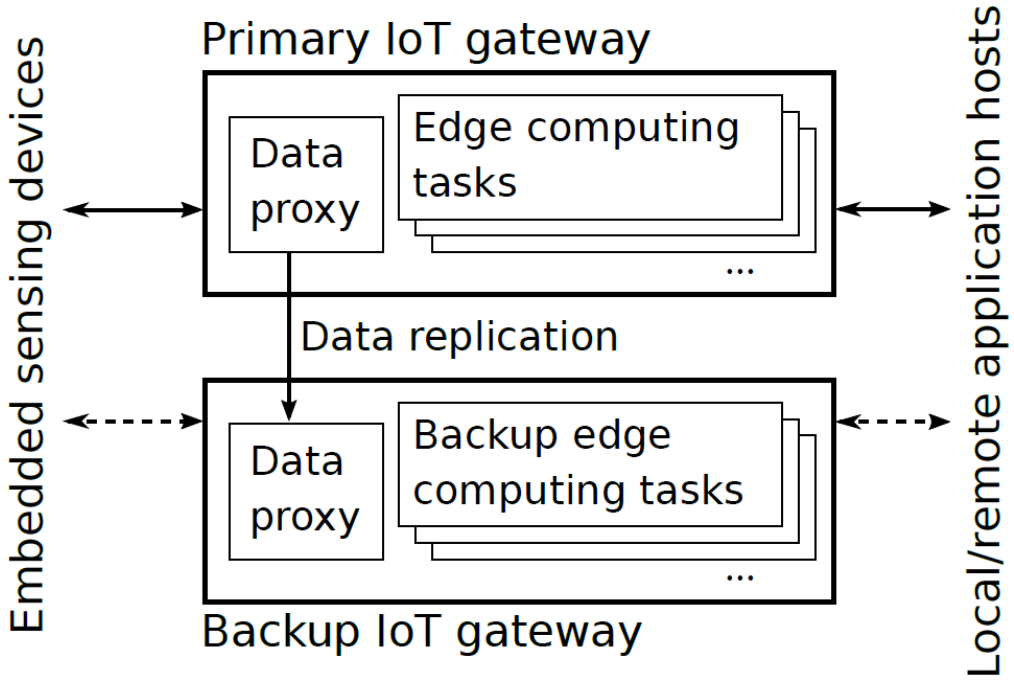
Specific IoT gateway requirements

- Quantitative requirements
  - Data subscriber cannot accept more than L_i consecutive losses for data topic *i*
  - Data subscriber imposes an end-to-end soft deadline for data
- Qualitative requirements
  - The gateway should not consume too much local network bandwidth



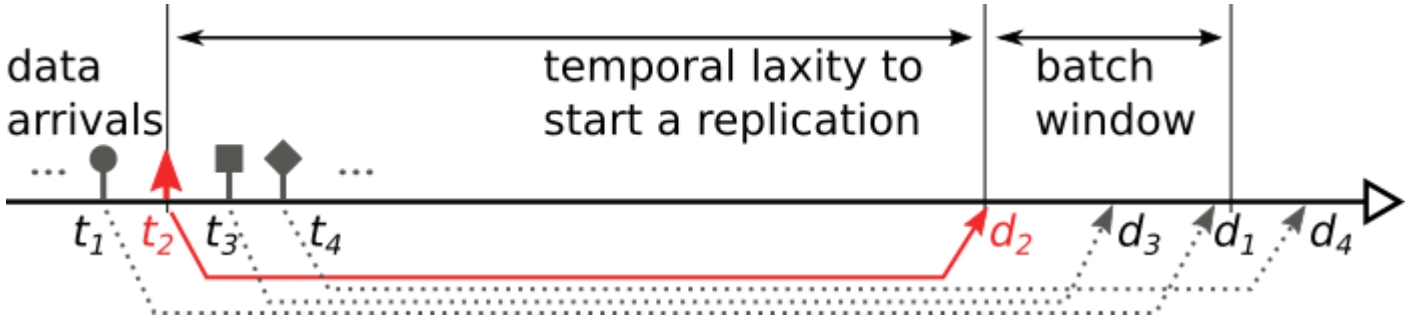
System model

- Publish-subscribe data model
  - with a minimum inter-publishing time for each topic
- Each embedded sensing device can only keep N_i latest data elements
- Primary-backup fault tolerance
  - consider the crash failure (fail-stop)
  - data replication to backup



Key idea for adaptive data replication

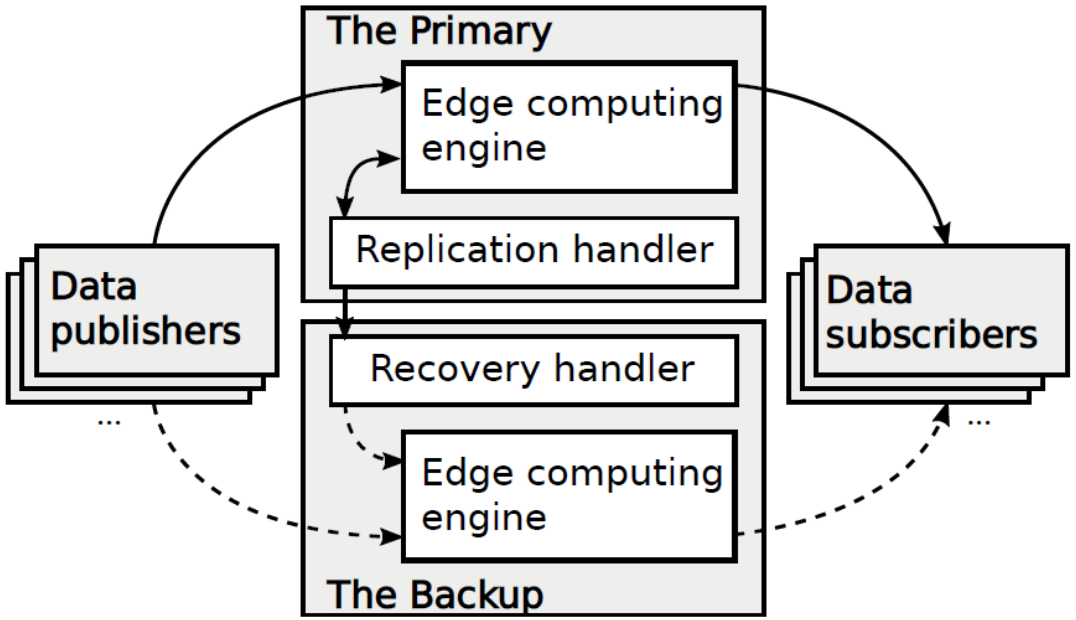
- In the IoT gateway, once data is processed/delivered, it is irrelevant
- Therefore, for each data, we may postpone replication activities to reduce the need of actually performing the replication!



  - Only start to replicate an arrival of data at the last starting time
  - Once start, perform a batch of pending replications

Adaptive data replication architecture

- Edge computing engine schedules both computing tasks and replication tasks using the EDF policy

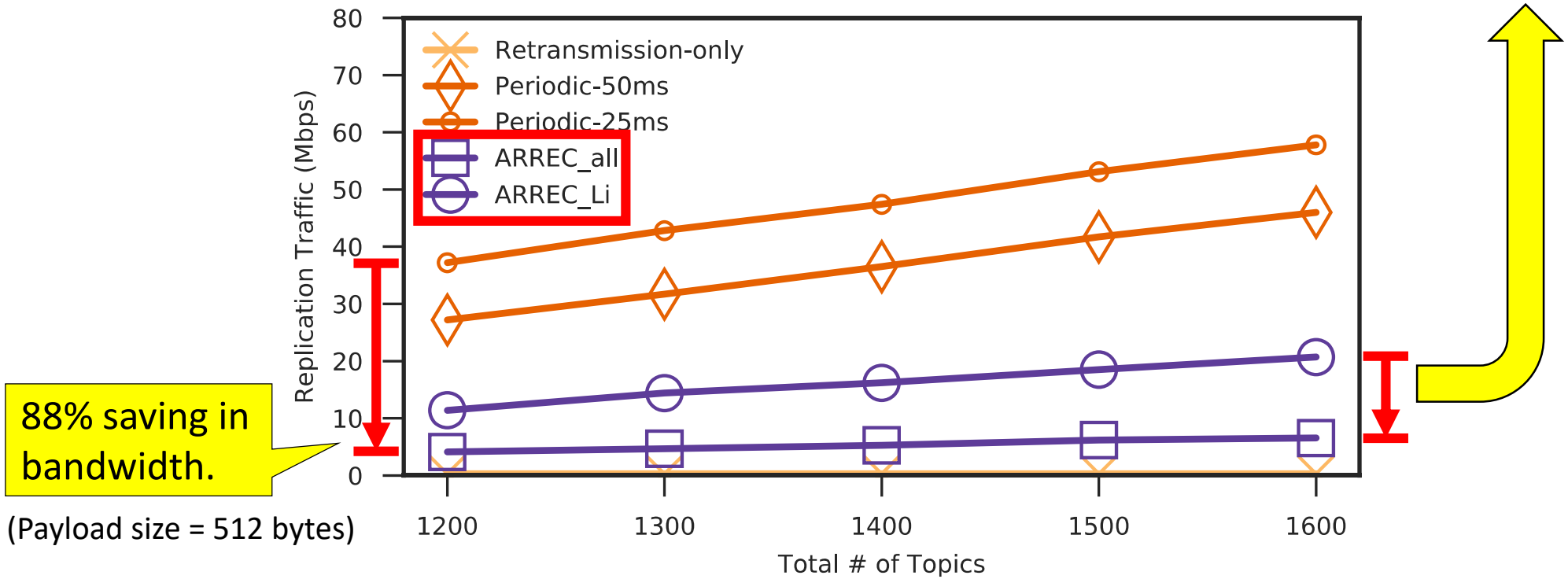


- Replication handler decides the intended rate of replication:

Lower intended rate -> tighter replication deadline
(see the paper for the analysis)

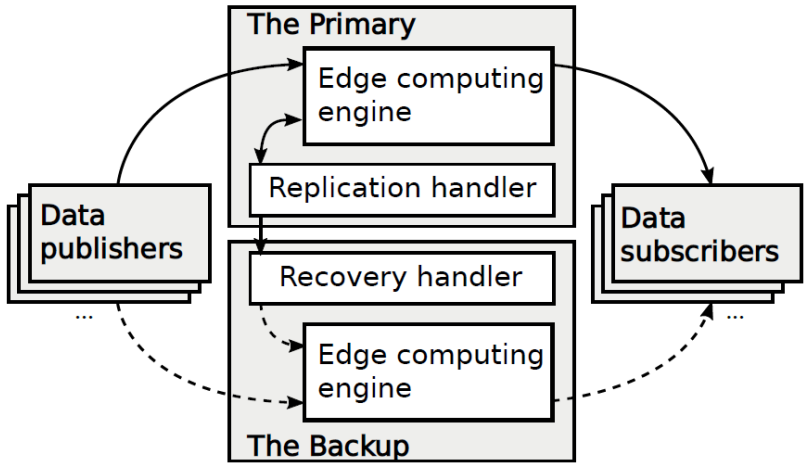
Empirical performance: efficiency in network bandwidth usage

A higher intended replication rate is preferred, because it permits a longer replication deadline, which in turn would allow the system to skip many more replication activities.



Concluding remarks

- Performant IoT edge computing is challenging
  - Reliability, timeliness, and resource constraints
- In this work we addressed the challenges for the case of IoT gateways
  - Key idea: adaptive data replication
- Our empirical validation shows that the proposed architecture can
  - meet the required levels of data-loss tolerance
  - save network bandwidth consumption
  - meet application-specific end-to-end deadlines



- Email us for further discussion: cw@ntnu.edu.tw