Supporting Efficient Workflow Deployment of Federated Learning Systems on the Computing Continuum

PhD Student: Cédric Prigent*, **Advisors:** Gabriel Antoniu*, Alexandru Costan*, Loïc Cudennec⁺

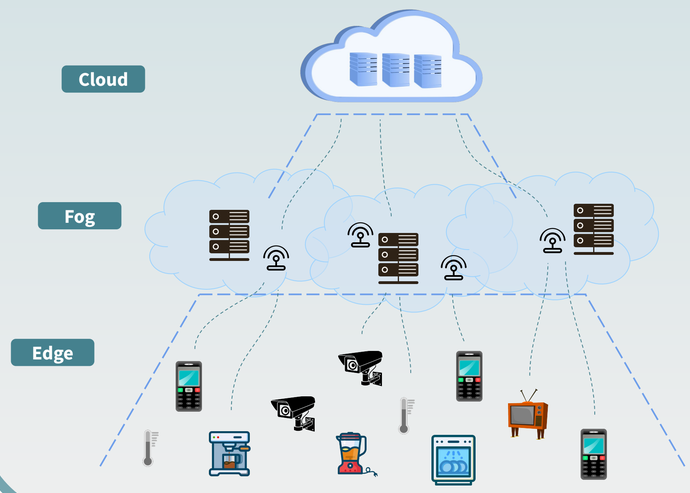
Contact: cedric.prigent@inria.fr

* University of Rennes, Inria, CNRS, IRISA - Rennes, France
⁺ DGA Maîtrise de l'Information - Rennes, France

Context

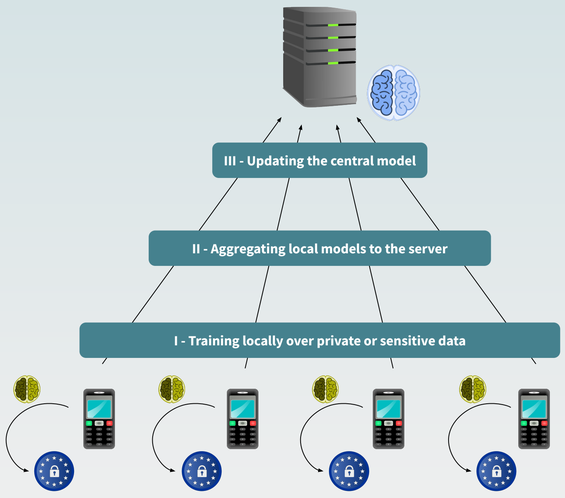
Computing Continuum

Applying complex workflows over Cloud-Fog-Edge Continuum



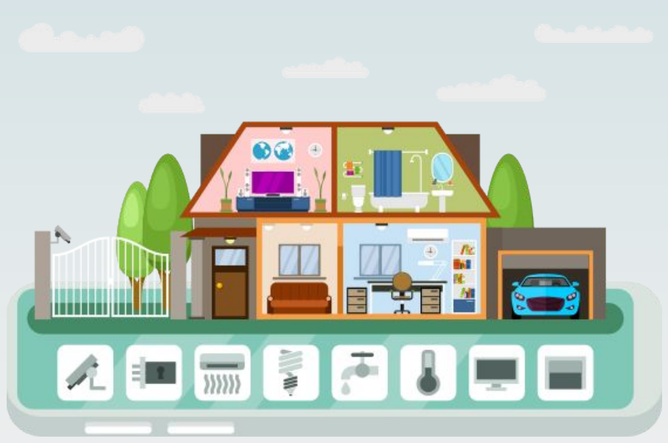
Federated Learning

Directly training over edge devices on sensitive data



Motivating Use Case

Applying FL to Smart Living using thousands of smart devices

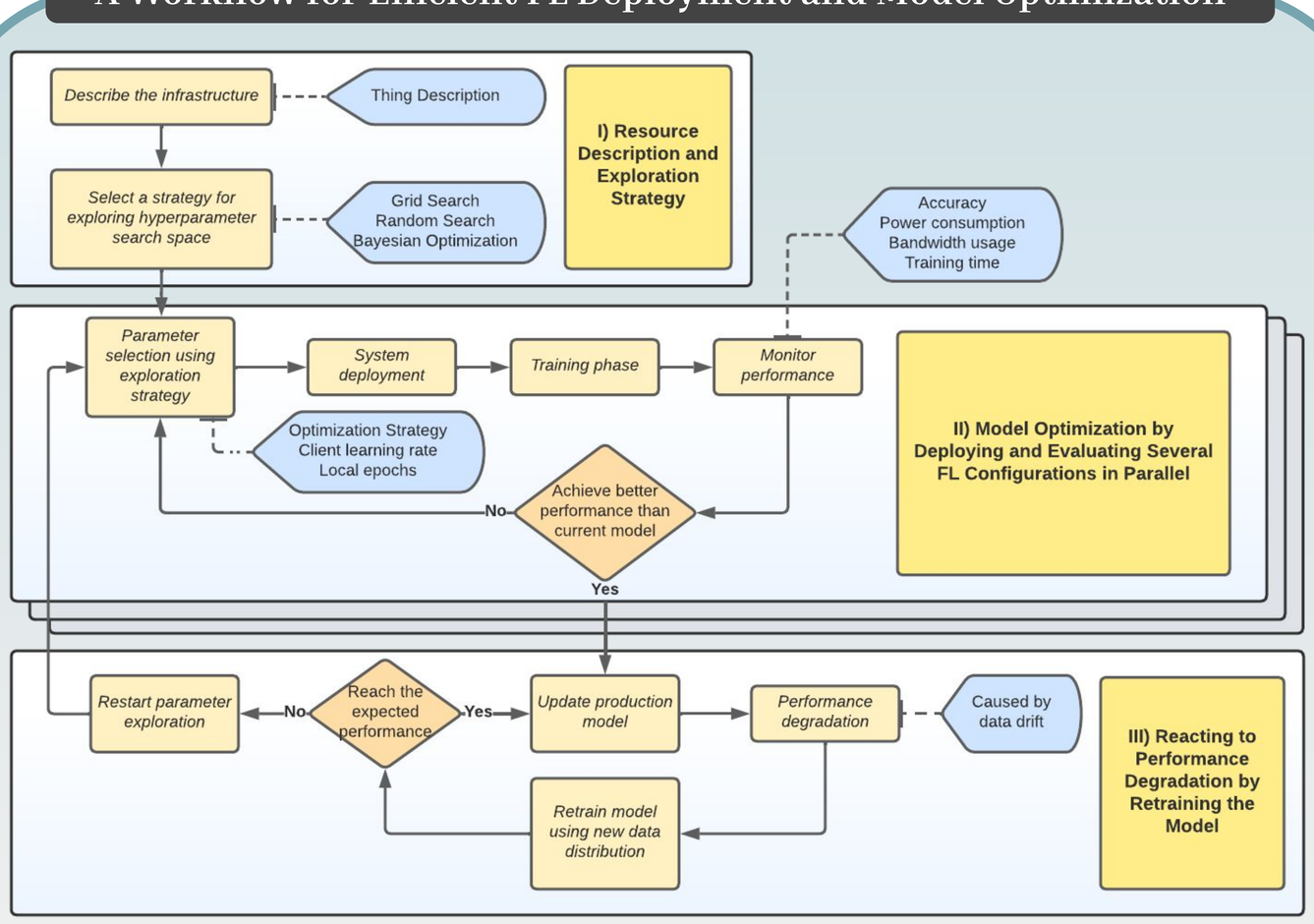


Problems

Heterogeneous devices	*IoT objects relying on different architectures require extra efforts to be supported in a same application (supporting different standards)*
Hyperparameter tuning	*Federated Learning introduces several additional hyperparameters in contrasts with standard ML, requiring deeper optimization*
Inconsistent performance	*Devices with very different usage result in data drift and inconsistent performance of the model among clients*

Our Proposal

A Workflow for Efficient FL Deployment and Model Optimization



How to improve interoperability across IoT platforms ?

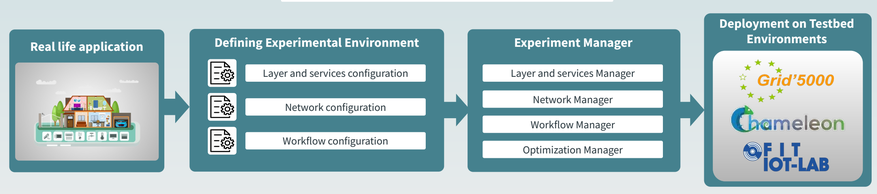
Thing Description is a W3C Standard for IoT platforms

Provide **entry points for IoT devices** by describing their interfaces

Allow IoT devices with different architectures to **communicate and interact with each other**

How to automatically deploy and monitor FL solutions ?

E2Clab Deployment Tool



Monitoring Metrics

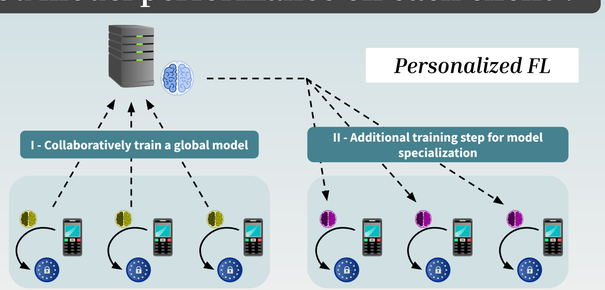
Parallel Deployment and Optimization

Reproducible Experiments

How to ensure good model performance on each client ?

Perform local adaptation step *on clients for model specialization*

Adapt models *for each clients with their local data distribution*



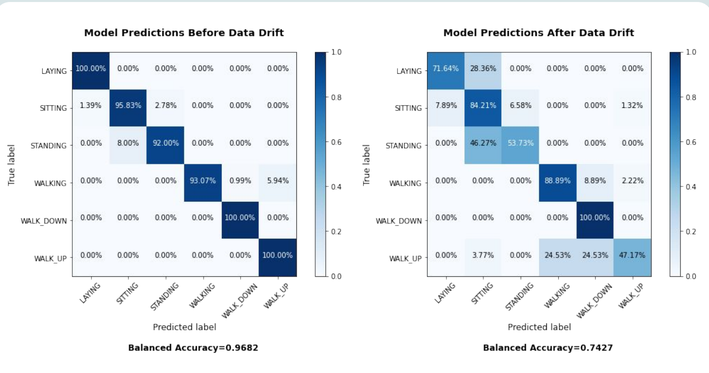
Early Results

Performing FL training over smartphone data from 8 users for recognition of Daily Living Activities

Impact of client learning rate and number of local epochs on training

GitLab repository

Impact of data drift on model performance



Takeaways

*We provided a solution to **automatically deploy and optimize FL workfloads in heterogeneous environments** using formal description of the underlying infrastructure, hyperparameter optimization and model retraining in case of data drift.*

Open Questions

*How to determine **which model is performing better** considering several **conflicting metrics** and clients with **different data distribution**? (e.g optimizing energy consumption, accuracy of the model and bandwidth usage)*

*Considering constrained-devices, **how often should models be trained** ? (minimizing battery usage of the device)*