Dynamic Intelligence Placement for Real-Time Data Analytics in Smart Buildings

Dimitrios Amaxilatis[§], Jhofre Ojeda[*], Kostas Ramantas[*], Themistoklis Sarantakos[§],
Nikolaos Tsironis[§], and Christos Verikoukis[†‡]
[*] *Iquadrat Informática, S.L.*, Barcelona, Spain
Email: {j.ojeda,kramantas}@iquadrat.com
[†] *University of Patras*, Patras, Greece
Email: cveri@ceid.upatras.gr
[‡] Industrial Systems Institute/Athena Research Center
[§] *Spark Works Ltd.*, Galway, Ireland
Email: {d.amaxilatis,tsarantakos,ntsironis}@sparkworks.net

***Abstract*—Smart building applications increasingly rely on machine learning for occupancy detection, but face a critical challenge: balancing edge computing's low latency and privacy benefits against cloud computing's elastic scalability. This demonstration presents a practical implementation of dynamic intelligence placement using the AC3 (Agile and Cognitive Cloud-edge Continuum) framework. Our system monitors environmental sensors (temperature, humidity, CO2) deployed in a 3D-printed room model to perform real-time occupancy classification, forecasting, and anomaly detection. When edge processing latency exceeds configured thresholds due to resource constraints or increased data rates, AC3 autonomously migrates the ML application pipeline from edge to cloud infrastructure using Kubernetes-based orchestration. Attendees will observe live migration events through interactive Grafana dashboards displaying real-time sensor data, ML inference results, resource utilization metrics, and migration triggers. The demonstration showcases how cognitive continuum management, integrated with data space principles from the 6G-DALI project, enables building operators to maintain both operational efficiency and occupant privacy by defaulting to edge execution while seamlessly leveraging cloud resources when necessary.**

***Index Terms*—Edge-Cloud Continuum, Adaptive Application Migration, Smart Building Analytics, ML Pipeline Orchestration, Privacy-Preserving Computing**

I. INTRODUCTION & MOTIVATION

Smart buildings equipped with IoT sensors generate continuous streams of data that enable critical applications such as energy-efficient HVAC control, space utilization optimization, and occupancy-aware security systems. These applications increasingly rely on machine learning models to detect occupancy patterns, forecast future states, and identify anomalies in real-time. However, deploying such ML-enabled applications presents a fundamental challenge: where should the intelligence be placed—at the edge for low latency and data privacy, or in the cloud for elastic scalability and computational power? Static deployment strategies fall short in dynamic environments. Cloud-only deployments impose unnecessary latency and raise privacy concerns by transmitting raw sensor data off-premises. Edge-only deployments, while privacy-preserving and responsive, struggle when sensor data rates increase or when multiple concurrent ML models exhaust local resources. The reality of modern smart buildings demands an adaptive approach that dynamically balances these trade-offs based on real-time system conditions [1], [2].

This demo presents a practical implementation of dynamic intelligence placement for smart building occupancy detection, leveraging the AC^3 (Agile and Cognitive Cloud-edge Continuum) framework. AC^3 is an EU Horizon project focused on efficient resource management across the cloud-edge continuum using AI-driven orchestration [3], [4]. Our system monitors environmental sensors (temperature, humidity, CO2) to perform occupancy classification, forecasting, and anomaly detection using ensemble methods and deep neural networks. When edge processing latency exceeds acceptable thresholds due to resource constraints or increased data rates, AC^3 autonomously migrates the application to cloud infrastructure, ensuring continuous operation while maintaining quality of service.

We demonstrate this capability through a tangible setup featuring a 3D-printed room equipped with real sensors (Figure 1), an edge computing device, and a live building monitoring interface. Attendees will observe real-time metrics including sensor readings, occupancy detection results, resource utilization, and latency measurements, witnessing autonomous migration events as system conditions change. This demonstration showcases how AC^3 enables building owners to maintain both operational efficiency and occupant privacy by keeping intelligence at the edge when possible, while seamlessly leveraging cloud resources when necessary.

This work is supported by the European Union's Horizon Program under the Agile and Cognitive Cloud edge Continuum management (AC3) project (Grant No. 101093129). This work has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101192750.

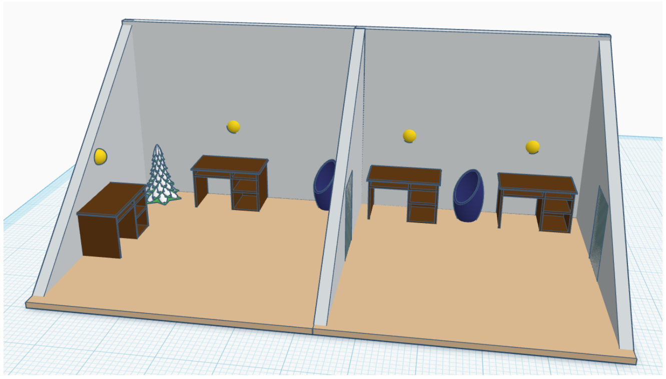


Fig. 1. The 3d office space used for the demo. Yellow dots represent sensors.

II. SYSTEM ARCHITECTURE

Our demonstration system implements a multi-layered architecture spanning the sensing, edge, cloud, and orchestration layers (Figure 2). A scaled 3D-printed room houses Sensirion sensors measuring temperature, humidity, and CO2—environmental parameters that serve as indirect occupancy indicators [5]. The edge node (a dedicated PC) hosts an ML pipeline with Random Forest classifiers for occupancy detection, LSTM networks for forecasting, and isolation forests for anomaly detection [6]. The same containerized pipeline deploys to cloud infrastructure for elastic capacity. AC^3's CECCM framework continuously monitors processing latency, triggering migration to cloud when thresholds (500ms per sample) are exceeded. The system integrates 6G-DALI data space principles for governance and MLOps workflows, ensuring model artifacts and sensor data remain accessible across deployment environments.

Data Management and MLOps Infrastructure: The system leverages data space principles from the 6G-DALI project to manage data sharing and governance across the edge-cloud continuum. Data spaces provide standardized interfaces for secure data exchange while maintaining data sovereignty, ensuring that sensor data and ML model artifacts are properly cataloged and accessible across deployment environments. The DataOps/MLOps pipeline handles model versioning, training data management, and automated deployment workflows, enabling seamless model updates and experiment tracking regardless of whether applications execute at the edge or in the cloud [7].

III. IMPLEMENTATION

The demonstration leverages the Cloud Edge Computing Continuum Management (CECCM) framework, a core AC^3 component that uses AI/ML and semantic context-awareness to predict resource constraints and manage application lifecycles across edge-cloud infrastructure. CECCM operates as a cognitive orchestration layer, continuously monitoring system state and autonomously triggering migrations when performance degrades beyond acceptable thresholds. This cognitive approach distinguishes our system from static rule-based orchestration, enabling proactive rather than reactive resource management [8].

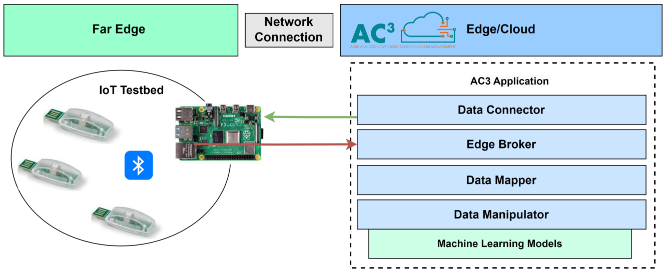


Fig. 2. Architecture of the demo

A. IoT Layer Setup

At the heart of the demo is the IoT layer, equipped with Sensirion sensors strategically placed throughout the Iquadrat office building to provide a real-time granular view of environmental conditions. They measure the following `environmental variables`: CO_2 concentration, temperature, and relative humidity.These variables are ideal for ensuring the health and comfort of the users at the office. Raspberry Pis serves as far-edge devices to collect and transmit data to processing layers.

B. Key Components of the AC3 Data Management System

The AC^3 Data Management System comprises three components: the Data Connector interfaces with Raspberry Pi far-edge devices to retrieve sensor data; the Data Mapper standardizes incoming data for downstream processing; and the Data Manipulator applies Isolation Forest and TensorFlow models for real-time anomaly detection and forecasting. RabbitMQ serves as the message broker, managing data transfer between far-edge, edge, and cloud [9]. Integration with 6G-DALI data space connectors enables federated data access while maintaining data sovereignty principles across the continuum.

C. Kubernetes-Based Deployment and Migration

The ML pipeline is containerized and deployed across two Kubernetes clusters (edge and cloud). Applications initially deploy to the edge cluster for minimal latency. CECCM monitors processing latency, CPU/memory utilization, and throughput. When latency exceeds 500ms per measurement, AC^3 orchestrates migration by deploying to cloud K8s, transferring model state, updating endpoints, and redirecting data flow. The Grafana dashboard automatically reconnects to the new endpoint, ensuring minimal service disruption.

IV. DEMONSTRATION

Our demo features interactive Grafana dashboards displaying real-time system performance, and a user-facing application designed for building managers and occupants. The main dashboard (Fig. 3) shows sensor data streams, data mapping latency, anomaly scores with processing times, predicted environmental conditions, and computational efficiency metrics.

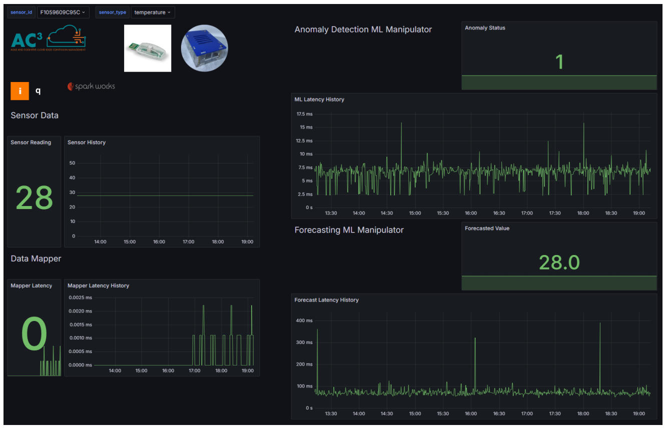
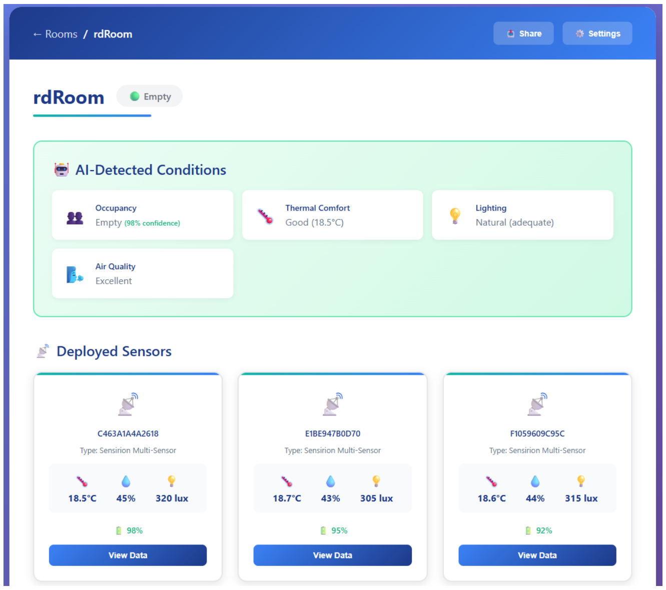


Fig. 3. Dashboard showcasing the raw and processed data flow.

A secondary dashboard monitors RabbitMQ operational metrics including message throughput, active connections, and resource utilization.



Occupancy
Empty (98% confidence)

Thermal Comfort
Good (18.5°C)

Lighting
Natural (adequate)

Air Quality
Excellent

C463A1A4A2618
Type: Sensirion Multi-Sensor
18.5°C 45% 320 lux
98%

E1BE947B0D70
Type: Sensirion Multi-Sensor
18.7°C 43% 305 lux
95%

F1059609C95C
Type: Sensirion Multi-Sensor
18.6°C 44% 315 lux
92%

Fig. 4. Our user-facing building management application.

The user-facing application offers an intuitive interface for building management (Fig. 4, displaying the current room occupancy status, environmental conditions (including temperature, humidity, and CO2 levels), and historical trends. Building managers can monitor multiple rooms simultaneously, receive alerts for anomalous conditions, and access occupancy forecasts for space planning and energy optimization. The interface remains fully functional throughout migration events, demonstrating transparent service continuity from the end-user perspective regardless of whether processing occurs at the edge or in the cloud.

Attendees witness live migration events triggered by simulated load increases: starting with baseline edge processing at low latency, we increase sensor sampling rates or enable additional ML models until latency exceeds the 500ms threshold, triggering AC^3's autonomous migration to cloud infrastructure. This cycle demonstrates the system's adaptive decision-making and can be repeated multiple times during the session.

We provide side-by-side comparison with cloud-only deployment: while cloud-only baselines show higher latency with all data transmitted off-premises, AC^3 maintains lower latency at the edge with data locality, migrating to cloud only when necessary—providing edge performance when possible with cloud scalability as backup.

V. CONCLUSION

This demonstration validates AC^3's cognitive continuum management for IoT applications, autonomously balancing edge responsiveness with cloud scalability. The system addresses persistent challenges in pervasive computing: resource-constrained devices, privacy-preserving processing, and adaptive architectures. By defaulting to privacy-preserving edge execution while automatically leveraging cloud resources under high load, AC^3 eliminates the binary deployment choice facing building operators and facility managers.

Our approach extends beyond smart buildings to industrial monitoring, smart cities, connected vehicles, and healthcare IoT where similar edge-cloud trade-offs exist. Ongoing work explores predictive migration strategies, multi-building hierarchical orchestration, and bidirectional migration policies. Integration with 6G-DALI data spaces enables secure multi-stakeholder scenarios for energy optimization and urban planning while maintaining data sovereignty. Conference attendees can observe multiple migration cycles and discuss applications in their own pervasive computing environments.

REFERENCES

[1] Y. Shi, K. Yang, T. Jiang, J. Zhang, and K. B. Letaief, "Communication-efficient edge ai: Algorithms and systems," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2167–2191, 2020.

[2] P. Bellavista, J. Berrocal, A. Corradi, S. K. Das, L. Foschini, and A. Zanni, "A survey on fog computing for the internet of things," *Pervasive and Mobile Computing*, vol. 52, pp. 71–99, 2019.

[3] L. M. Vaquero and L. Rodero-Merino, "Finding your way in the fog: Towards a comprehensive definition of fog computing," *SIGCOMM Comput. Commun. Rev.*, vol. 44, p. 27–32, Oct. 2014.

[4] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5g network edge cloud architecture and orchestration," *Commun. Surveys Tuts.*, vol. 19, p. 1657–1681, July 2017.

[5] S. Yi, C. Li, and Q. Li, "A survey of fog computing: Concepts, applications and issues," in *Proceedings of the 2015 Workshop on Mobile Big Data*, Mobidata '15, (New York, NY, USA), p. 37–42, Association for Computing Machinery, 2015.

[6] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, 2017.

[7] 6G-DALI Consortium, "6g-dali: Data-driven application lifecycle for 6g." https://6g-dali.eu, 2024. EU Horizon Project, Grant No. 101096466.

[8] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, and Y. Agarwal, "Sentinel: occupancy based hvac actuation using existing wifi infrastructure within commercial buildings," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, SenSys '13, (New York, NY, USA), Association for Computing Machinery, 2013.

[9] R. G. Hegde, "Low latency message brokers," in *International Research Journal of Engineering and Technology (IRJET)*, 2020.