Mediterranean Smart Cities Conference (MSCC 2024)
2-4 May, 2024, Martil, Morocco

Cooperative and Connected Mobility Services in the Cloud-Edge Continuum with Function As A Service Technology and AI-enabled Orchestration

Antonio Lalaguna
ACISA
Barcelona, Spain
antonio.lalaguna@acisa.es
0009-0003-4749-1695

Paul Townend
Umeå University
Umeå, Sweden
paul.townend@umu.se
0000-0002-9698-8361

Behnam Ojaghi
ACISA
Barcelona, Spain
behnam.ojaghi@acisa.es
0000-0002-5642-0974

Constantino Vázquez
OpenNebula Systems,
Madrid, Spain
cvazquez@opennebula.io
0000-0002-3059-5063

Abstract—**We propose a novelty system to manage Traffic Priority at city intersections by means of our Mobility-Hub (M-Hub), a next-generation Traffic Light Controller that leverages the power of cloud-edge continuum computing, Digital Twin, and Cellular Vehicle-to-Everything (C-V2X) technologies to transform traffic management into a dynamic and intelligent system. M-Hub acts as an open-edge computing platform, enabling real-time data processing, 3rd parties containerized applications and decision-making at the network edge. COGNIT is an open-source cloud-edge continuum framework, that offers many improvements for next-generation Intelligent Transportation Systems (ITS). The continuum allows for the integration of diverse data sources, including vehicular data from C-V2X communication, real-time traffic information from detectors or cameras, and other environmental data, to seamlessly generate Digital Twins in the ACISA smart mobility platform, SATURNO. By combining this data with advanced traffic optimization algorithms implemented in the COGNIT infrastructure, M-Hub can dynamically adjust traffic signal timings, optimize traffic flow, and reduce congestion with optimal use of computational resources.**

M-Hub has the potential to revolutionize urban mobility, enhancing safety, improving efficiency, and reducing environmental impact.

***Keywords*— Cloud-Edge Continuum, Digital Twin, 5G, C-V2X, Urban Mobility, Intelligent Transport Systems.**

I. INTRODUCTION

Increasing urbanization and growing demand for mobility have placed significant pressure on existing traffic management systems, where urban mobility is facing unprecedented challenges due to increasing traffic volumes, expanding urban areas, and the rise of autonomous vehicles and smart cities. Traditional traffic light controllers, often based on static and fixed signal timings and outdated technology, are limited in their capabilities and struggle to keep pace with these changes and the complexities of modern urban environments. In urban environments, many ITS sub-systems (e.g., traffic light controllers, smart cameras, inductive loop detectors, Bluetooth sensors, etc.) must coexist to monitor and control traffic in situations of high complexity.

Nowadays, the emergence of new sources of vehicular data through Vehicle-to-everything (V2X) communication has reactivated the need for new solutions, as most existing systems cannot address the *Quality of Service* (QoS) requirements of traffic demands. V2X networks consisting of *Vehicle-to-vehicle* (V2V), *Vehicle-to-pedestrian* (V2P), *Vehicle-to-Infrastructure* (V2I), and *Vehicle-to-Network* (V2N) communication are envisioned as a fundamental enabler for ITS, where vehicles are equipped with an *On-Board Unit* (OBU), which includes a V2X communication stack and allows on-board sensor data to interact with neighbouring vehicles, *Road Side Units* (RSUs), and cloud applications over cellular connectivity [1][2].

In principle, V2X connectivity enables vehicles to perceive beyond their onboard sensors and realize *Connected Cooperative and Automated Mobility* services (CCAM). These services are intended to improve traffic efficiency and road safety for ITS. To improve transportation efficiency, safety, and comfort on the road, V2X services have various and complex application scenarios, ranging from enhancing real-time safety and efficiency on the road to autonomous driving. Moreover, it is believed that the number of autonomous cars will increase significantly owing to urbanization and technological progress, enabling massive usage of V2X communications and promoting the development of intelligent autonomous vehicles. This increasing demand poses significant new communication issues in V2X networks [3] - [5].

As a result, existing LTE/5G networks will face significant capacity limitations, resulting in new scientific and technical challenges for V2X communication networks in terms of data rate, latency, coverage, intelligence level, network connectivity, vulnerabilities, and security issues [3]. *To help smart-city service providers manage this complexity and deploy faster to meet the QoS requirements of different services, platforms need to be interoperable, and support applications that can be deployed either in Edge, Cloud, or On-Premises, and can be managed remotely using Cloud-native practices*.

Cloud computing infrastructures have evolved from centralised data centres to distributed federations of processing and storage resources – often known as the "Computing Continuum". The continuum [6] offers many opportunities for next-generation smart transport systems, including lower network latencies, better scalability, and smart offloading of computation on resource-constrained devices – but at the cost of greatly increased complexity in development, deployment, and management. COGNIT aims to resolve this by facilitating the orchestration of distributed and heterogeneous resources and abstract complexity to final clients (e.g. M-Hub).

Besides the infrastructure complexity, many intersections in a big city are equipped with a variety of information systems from different providers that add context information to the traffic operators, and eventually to road users. We are modelling city intersections and creating its Digital Twin (DT) with the data provided by all those systems, to present a unified interface to traffic operators in the *Control Center*, enabling the development of predictive traffic models and simulations to increase safety, reduce emissions and facilitate interaction with users and DTs in other intersections.

Every intersection within a city varies in terms of its topology, traffic model, M-Hub program, and the specific manoeuvres that a vehicle may request priority for, resulting in a high degree of complexity. Each intersection has its counterpart digital representation, a.k.a. Digital Shadow, in Saturno, holding updated historical data about its layout, current traffic status, subsystems status and alarms (M-Hub, detectors, others), etc. By incorporating the traffic simulator, the Digital Shadow evolves to a Digital Twin, who, will act over the intersection traffic lights to grant or deny the priority based not only on current traffic status, but also on its simulated predictions.

The remainder of this paper is organized as follows: Section II introduces how ACISA address the challenge encountered by city traffic operators dealing with multiple providers, applications, and subsystems, through the establishment of an open edge-cloud infrastructure. Next, the objectives of this approach are presented in Section III. The methodology followed is explained in Section IV, and Section V concludes the paper.

This paper presents its solution over a framework being developed in the EU project COGNIT [7], a 6M Euro project funded by the EU Horizon Europe program[1] which aims to build an open-source software stack that enables the deployment of *Function-as-a-Service* (FaaS) technology in the continuum, with the aim of artificial intelligence techniques to optimally orchestrate available computing resources.

II. CONTRIBUTION

Over the last decades, new technologies have been incorporated into road intersections to enrich context knowledge (intelligent cameras, Bluetooth, and wireless sensors, V2X RSUs, pollution and noise sensors, dynamic information panels, *Vulnerable Road Users* (VRU) detection, etc.) This further complicates the work of operators in traffic control centers, where a plethora of applications from different providers coexist in a non-integrated fashion.

To address the challenges mentioned above, we propose our advanced Traffic Light Controller (TLC), named *Mobility-Hub (M-Hub) [8]*, which is designed to bring together the core functional elements of cloud-edge platforms for smart mobility solutions through containerized and intelligent applications (Fig. 1). These are developed and implemented within M-Hub itself, which enables complex or simple services with different QoS requirements to be managed at the edge (*i.e.*, M-Hub) utilising unified interfaces and data formats.

Figure 1. M-Hub: ACISA's Next Generation Traffic Light Controller

Considering that urban traffic infrastructures are owned by the city councils, enabling an open, standard-based far-edge platform will allow cities to offer a robust infrastructure 3^{rd} parties providers, reutilizing existing traffic infrastructure. Some of the key features of M-Hub are: i) providing an environment to virtualize, communicate and deploy advanced computing services

[1] This work was supported by the European Commission through the Horizon Europe project SovereignEdge.Cognit under Grant Agreement 101092711.

and applications, and to make them available to cities and other service providers; ii) providing a local gateway and secured communication hub between any subsystems deployed at the intersection and the cloud continuum; iii) providing an infrastructure prepared for critical services that require low latency, high bandwidth and local communication with other subsystems (such as CCAM applications or safety applications based on intelligent video processing).

M-Hub provides the interface with the cloud-edge continuum platform to other subsystems installed at road intersections by different providers. Besides applications developed by ACISA, it also allows third parties to develop applications that may interact with the different subsystems in the road intersection to enrich their own data and algorithms, by means of data contract agreements. M-Hub provides local data processing *i.e.*, data quality assessment, missing data detection, time series inference predictive models and data protocol transformation to comply with standards.

M-Hubs are normally connected to our smart mobility platform named SATURNO [9], a UNE 178104:2017 compliant platform, built as a cloud-native *Software as a Service* (SaaS). SATURNO is also an open suite with a powerful User Interface (UI) and Application Programming Interface (API) that enables an integral management of mobility systems such as urban traffic, low emissions zones, access control, environmental monitoring, administrative back-office integration, etc.

For applications that requires low latency, high availability, and higher computational power than what an M-Hub can offer, additional computing resources are allocated in the Control Centre of the city in a 3-tier architecture. Based on latency requirements, storage needs, and/or legal agreements, SATURNO services may be deployed either in the cloud, in an on-premises data center or as a hybrid deployment.

To facilitate the management of those distributed resources, we are participating in the COGNIT project to incorporate an advanced AI-enabled *serverless* framework as part of our infrastructure. Serverless Computing [10] allows application developers to upload individual *code functions* to a Cloud or Edge system; this system is then itself responsible managing the placement, scaling, fault-tolerance, etc. of these functions, and abstracts infrastructure complexity to developers.

The COGNIT framework (shown in Fig. 2) *efficiently manages distributed resources as a continuum infrastructure*. Its serverless paradigm approach not only simplifies the setup and maintenance operations but also significantly reduces operational costs, using Artificial Intelligence (AI) models to place and manage functions optimally.

At the heart of COGNIT, application developers and device clients submit functions to a Provisioning Engine, which provides a Serverless runtime for the function. The Provisioning Engine then interacts with a Cloud-Edge manager, which has the goal of *placing the runtime on an appropriate node within the Continuum*.

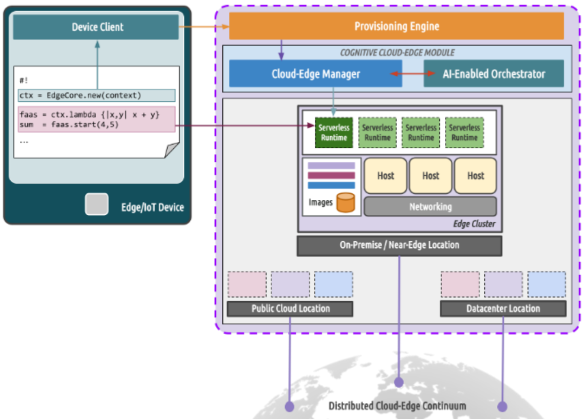

Figure 2. The COGNIT Computing Continuum architecture

Moving computation to the edge poses new security risks. Edge Computing needs intelligent security, like privacy-preserving methods such as federated learning and advanced authentication, to ensure low latency and resilient anomaly detection and remediation.

The adoption of a serverless cloud-edge continuum paradigm represents a novelty itself, but its application in in the traffic management domain disrupts the conventional landscape and enables new possibilities for innovations in a sector traditionally controlled by proprietary solutions.

III. OBJECTIVES

The main objectives of the solution proposed are i) to demonstrate the benefits of using the cloud-edge continuum architecture described in Sec. II to develop more secure, available, lower latency smart mobility services, such as Traffic/Transit Signal Priority; ii) to support an innovative new serverless paradigm for edge application management and enhanced digital sovereignty for users and developers; iii) to enable on-demand high available & low latency deployment of applications using on-premises city infrastructure (increasing resource utilization); iv) to optimize data placement according to changes in energy efficiency heuristics and application demands and behaviour; v) to enable secure and trusted execution of the serverless runtimes that support the new V2X services that ACISA installed in the city; vi) to achieve a faster pace of innovation, more resilient systems, and less proprietary technologies; and vii) to align with the CCAM objective to create a more user-centred and inclusive mobility system, increasing road safety while reducing congestion and environmental footprint.

Using the cloud-edge continuum infrastructure described in Sec. II, we aim to leverage state of the art technologies in the cloud and edge computing fields that will benefit our customers by providing functionalities only available on public cloud providers nowadays. *To this aim, we leverage the privileged location held by each TLC in the layout of a city to provision an enhanced distributed computing capacity, which enables a low-latency high-capacity infrastructure to deploy advanced services for traffic management*.

Given that the overall physical infrastructure is owned by city councils, providing such open, standard-based infrastructure will allow cities to deploy robust smart-city

IoT & mobility services relying on its already deployed traffic infrastructure.

IV. Methodology

M-Hub enables low latency Infrastructure to Vehicle (I2V) services – based on ETSI and 3GPP standards – such as Traffic/Transit Signal Priority (TSP), Time-To-Green information (TTG) and GLOSA (Green Light Optimal Speed Advisory) as harmonized in C-ROADS project. These services aim to enhance driving experience by providing speed advice, traffic light information and countdowns or green priority, to reduce energy consumption and minimize stops, promoting eco-friendlier and energy-efficient driving. They rely on standardized messages such as Cooperative Awareness Messages (CAM), Decentralized Environmental Notification Messages (DENM) and In-Vehicle Information Messages (IVIM) in accordance with the standard ETSI TS 103 301 [11].

Those M-Hub capabilities are being used to implement the TSP service in Granada, that aims to adapt the status of the traffic lights for a prioritized vehicle, such as a public transport. These vehicles may – for example - receive earlier green signals, enabling them to move quickly, shorten travel times, and increase road safety. This is achieved by modifying TLC phase timings, when necessary, either by extending the green phase, reducing the red one, or even forcing a new phase (in the case of emergency vehicles). The precise and detailed information that can be obtained with the aim of V2X systems, and the additional computational power at the intersection provided by the M-Hub, allows us to be more precise, and consider additional information before granting priority. Additional information may include traffic intensity and occupation, incidents, traffic jam, near bus stop, weather conditions or pollution, together with contextual information the public transport might send (*i.e.*, bus location, speed, direction, delay or advance with respect to the schedule, passenger occupancy, etc.)

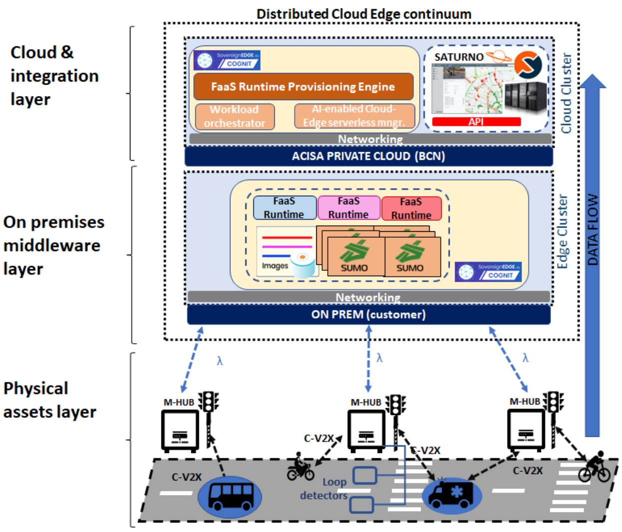

Figure 3. C-V2X based Bus priority use case within a cloud-edge Continuum infrastructure.

M-Hub collects traffic information from the subsystems it connects continuously, to synchronize with the DT of the road intersection. Periodically it may also need to perform simulations to evaluate future traffic status and *What-If* scenarios based on current context data. These is simulation tasks are off-loaded to the COGNIT framework, which will provision the resources within the most optimal node available in terms of latency, energy, or computational capacity (Fig. 3).

Upon receiving a priority request from an authorized vehicle, the M-Hub may initiate a function call to the COGNIT FaaS service including additional contextual information to execute on demand traffic simulation by the Sumo tool. This is some heavyweight functionality that wouldn't make sense to deploy on the client and will be leveraged using the remote FaaS.

The continuum infrastructure is held by servers allocated either in the public cloud or at ACISA's private cloud data center, along with on city council premises servers, that may be used to store privacy sensible data, video, or city council's owned licensed software.

ACISA deployed a Traffic Signal Priority service based on C-V2X communication technology in Granada, a city in southern Spain. It is based on direct interaction between the C-V2X subsystem and ACISA's M-HUB in the road intersections, edge resources at Center for Comprehensive Mobility Management of Granada [12], and on SATURNO deployed in our private cloud in a 3-tier architecture as shown in Fig. 3.

In this real scenario, we have deployed a Proof of Concept (PoC) where we use the COGNIT framework to offload the launch of traffic simulations of an intersection under interest. We follow the standard UNE-CEN ISO/TS 19091 that defines the use V2X communications for applications related to signalized intersections.

A public bus requests priority by sending a standardized V2X Signal Request Extended Message (SREM), once it detects broadcasted messages with intersection's layout and TLC phase information from M-Hub through Map Extended Message (MAPEM), and Signal Phase and Timing Extended Message (SPATEM) respectively, as defined in ETSI TS-103-301 and SAE J2735 [13]. The SREM includes essential details such as the bus ID, Lane ID, and, when available, passenger occupancy level.

M-Hub processes V2X messages, and filters those requesting for traffic light priority. Every time an SREM is received, it first evaluates if it is a valid request; then, it will launch an internal process to locally evaluate the permission. M-Hub tracks bus positions by means of Cooperative Awareness Messages (CAM) sent by the vehicle every 100ms, to ensure that buses do not stop unexpectedly due to traffic congestion, bus stops, or any other unplanned reason. Once the bus reaches a specified location (particular to each lane of the road intersection, at each intersection) the final evaluation takes place: M-Hub evaluates current traffic status based on feedback from the Digital Twin, verifies any delay/advance of the bus with respect to the planned schedule, evaluates the last simulation results, and finally initiates the necessary traffic light phase changes to ensure it will find a green light on time (Fig. 4).

Every interaction with the COGNIT framework is performed by means of lambda functions executed on the serverless COGNIT infrastructure. Should the M-Hub need extra data from external systems, it can request the missing information with simple lambda functions, abstracting the M-Hub from the complexities in the backend where information may come from either SATURNO suite, *Automatic Vehicle Location* (AVL) systems from public transport operator, city council information endpoints, public administrations databases or any other related information system.

This abstraction is particularly important for M-Hub to become an open platform for third parties. These external parties should not have access to internal infrastructure details, to ensure security and trust, simplicity of development and deployment, decoupling, and low-footprint applications.

Up until today, all the necessary information to process traffic priority requests is held in centralized systems like SATURNO, but the aim is to move these decisions closer to the requestor through the COGNIT serverless framework.

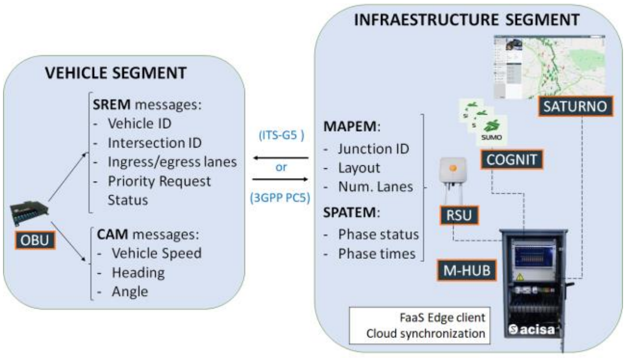

Figure 4. Main ETSI & J2735 messages used in TSP Service

As explained in Section II, both local-based and Control Center-based approaches present their own challenges. Increasing the computation power of a TLC, enables it to take local decisions that decrease response time while ensuring service availability. Authors in [14] present a model for efficient multi-player and multi-task computation offloading in VR applications, leveraging mobile edge-cloud computing to ensure optimal network latency and energy consumption. They propose an architecture where main computational tasks are shared between user-side VR devices, and telco edge resources (in this case MEC servers instead of the city council's on-premises) and demonstrated a reduction in network latency and energy consumption between 27% and 11% when compared to cloud, edge or user device execution alone. We are aiming for even further improvements as we rely on edge servers that were already present in the control center, fiber optic connection between far edge M-Hubs and continuum infrastructure, and the COGNIT capabilities to optimize computation resources in terms of latency and energy consumption.

V. Conclusions

M-Hub represents a significant step forward in the evolution of traffic management. By leveraging cloud-edge continuum computing, Digital Twins, and C-V2X technology, M-Hub enables a more intelligent, adaptive, and data-driven approach to mobility management. This is particularly effective when the COGNIT Serverless framework is used for important smart city use cases such as public bus priority. With M-Hub's ability, combined with the SATURNO cloud platform, it is possible to optimize traffic flow, reduce congestion, and enhance safety, giving M-Hub the potential to transform the way we move around our cities. As we move towards a future of smart cities and autonomous vehicles, M-Hub will play a critical role in ensuring the efficient, sustainable, and safe movement of people and goods in our urban environments.

The use case presented provides performance information about i) applying V2X technology for public bus and emergency vehicle prioritization; and ii) using a serverless framework to dynamically provision resources to execute heavy or centralized tasks. This combination presents a novelty in the smart city context. In ongoing and future work, the Digital Twin of each intersection will be completed where TLC, road layout, traffic and V2X message propagation and collisions will be modelled for urban scenarios and will then be generalized to be used in different cities and interurban scenarios.

The proposed solution is not devoid of risks and potential challenges. One significant limitation arises from the requirement for highly skilled engineers capable of implementing sophisticated solutions within the traffic management domain. The concept of the cloud-edge continuum offers a novel approach to extending cloud-native tools to the edge domain, thereby unifying resources across both environments. However, the heterogeneous nature of edge nodes presents distinct challenges compared to their cloud counterparts. Access control, cybersecurity [15], communication protocols, licensed applications, and other IT governance issues, pose significant challenges in achieving efficient management within this paradigm.

COGNIT is situated under the general framework of the European Cloud Edge IoT continuum initiative [16]. This initiative is dedicated to facilitating the comprehension and advancement of the Cloud, Edge, and IoT (CEI) Continuum, fostering collaboration among different research projects, developers, suppliers, business users, and prospective adopters, to achieve a widespread adoption. Other consortia tackling analogous challenges to COGNIT include CODECO [17], FLUIDOS [18], or AC3 [19].

References

[1] S. Chen et al., "Vehicle-to-everything (v2x) services supported by LTE-based systems and 5G," IEEE Commun. Standards Mag., vol.1, no. 2, pp. 70–76, Jun. 2017.

[2] "Evolutionary V2X technologies toward the internet of vehicles: Challenges and opportunities," Proc. IEEE, vol. 108, no. 2, pp. 308–323, Feb. 2020.

[3] W. Saad, et al., "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," in *IEEE Network*, vol. 34, no. 3, pp. 134-142, May/June 2020, doi: 10.1109/MNET.001.1900287.

[4] S. Nayak, et al. “6G communications: A vision on the potential applications,” 2020, arXiv:2005.07531.

[5] I. F. Akyildiz, et al. “6G and beyond: The future of wireless communications systems”, IEEE Access, vol. 8, pp. 133995–134030, 2020.M. Muhammad and G. A. Safdar, “Survey on existing authentication issues for cellular-assisted V2X communication,” Veh. Commun., vol. 12, pp. 50–65, Apr. 2018.

[6] A. Aske, X. Zhao, "Supporting Multi-Provider Serverless Computing on the Edge", in Workshop Proc. of 47th ACM Int. Conf. on Parallel Processing, 2018

[7] P. Townend et al. “COGNIT: Challenges and Vision for a Serverless and Multi-Provider Cognitive Cloud-Edge Continuum.” In Proceedings IEEE International Conference on Edge Computing and Communications, Chicago, IL, USA, 2023, pp. 12-22, doi: 10.1109/EDGE60047.2023.00015

[8] Available at the ACISA website: https://www.acisa.es/en/urban-traffic-technology/#mhub

[9] Available on the ACISA website: https://www.ACISA.es/innovacion/saturno/

[10] E. Jonas et al. "Cloud programming simplified: A berkeley view on serverless computing.", arXiv preprint arXiv:1902.03383

[11] Available at: https://www.etsi.org

[12] Available at:

[13] http://www.movilidadgranada.com/tra_cgim.php?idioma=en

[14] Available at: J2735_202309: V2X Communications Message Set Dictionary - SAE International

[15] Alshahrani, Abdullah et al., (2020). “Efficient Multi-Player Computation Offloading for VR Edge-Cloud Computing Systems” Applied Sciences. 10. 5515. 10.3390/app10165

[16] Marin, E., Perino, D. & Di Pietro, R. “Serverless computing: a security perspective.”, J Cloud Comp 11, 69 (2022). https://doi.org/10.1186/s13677-022-00347-w

[17] Available at: https://eucloudedgeiot.eu/

[18] G. Koukis et al. “An Open-Source Experimentation Framework for the Edge Cloud Continuum”, in INFOCOM 2024 CNERT: Computer and Networking Experimental Research using Testbeds Workshop (CNERT), Vancouver, Canada, mar. 2024. doi: 10.5281/zenodo.10840008.

[19] M. Iorio et al. "Computing Without Borders: The Way Towards Liquid Computing," in IEEE Transactions on Cloud Computing, vol. 11, no. 3, pp. 2820-2838, 1 July-Sept. 2023, doi: 10.1109/TCC.2022.3229163.

[20] C. Symvoulidis et al. "Dynamic deployment prediction and configuration in hybrid cloud / edge computing environments using influence-based learning," 2023 10th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Palembang, Indonesia, 2023, pp.315-320,doi: 10.1109/EECSI59885.2023.10295768.