UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Expression of Interest on future visions and research directions 2025-27 in the area of Cloud-to-Edge-to-IoT for European Data

This document expresses the view of two research groups: UPV's SATRD and SRIPAS, which are currently leading the projects H2020 ICT-56 ASSIST-IoT and HE DATA-01-05-2021 aerOS. UPV coordinates both RIAs and SRIPAS acts as Technical Coordinator in the former and as key technological contributor and prominent scientific dissemination agent in the latter. Their position in the current research landscape, and informed perspectives have motivated this Expression of Interest. The content below summarises the view of these two groups but does not necessarily represent the will of the whole ASSIST-IoT's or aerOS' Consortiums, which are formed by heterogeneous partners with their corresponding interests.

Motivation

One of the priorities of the EC for 2030 is a successful digital transformation of Europe for people and businesses. Here, the "fit for the digital age" dictates that cloud and edge computing are essential to drive a sustainable digital transformation. In cloud computing, US holds the lead (Europe is below 20% of the market share), but reports like H-CLOUD's whitepaper, EAIDCE, FCC and EAA&BI coincide in pointing to the edge-to-cloud hybrid paradigm (thus, the continuum) as a strategic technology towards European leadership in the digital space (aligned with EU Data Strategy). In addition, Europe is taking a leading role in the ethical and regulatory fields following the release of the AI act, in contrast to the few regulatory steps applied by US or the controversial Chinese and Russian approaches. Furthermore, European values, research capabilities and infrastructure provide the tools for this competitiveness to take place. Envisioning the research priorities that will lead the continuum (embedding cloud, edge, IoT and everything in between) is crucial to properly position our continent. Hence, it can be stipulated that, in edge-cloud computing continuum, a whole new market will open up that should be (by values, data protection and perspective) dominated by EU.

Besides, amid a chip shortage and a continental technological dispute, Europe must invest in its own resources, reducing dependence on specific supply chains, ensuring resilience during disruptive events.

Moreover, Europe should be aware of any room to innovate. Nowadays, the continuum (i) is complicated, (ii) heterogeneity remains unresolved, (iii) applications depend heavily on underlying libraries, CPU architectures, and resource management frameworks, (iv) existing technologies do not fit together seamlessly, (v) applications are unnecessarily heavy, due to bloated container images and ineffective use of available resources, (vi) there is reluctance to share data and (vii) trustworthiness concerns arise when using edge computing. Besides, current mechanisms and solutions are not prepared for scalability and adaptability, requiring deep and complex updates whenever new innovations come out.

Current status

Several projects have tackled relevant research challenges in recent times (more than 35 projects gathered under EUCEI CSA management), but vision must now shift to what comes after the currently developing continuum-creation technologies. Here, the advance speed is outstanding. In this sense, formal research actions must go beyond pure scientific literature and results analysis (risking outdated results) and must broaden their scope to other sources such as open-source development projects or successful use-cases in key vertical industries, but extendible to any other. Indeed, agile, dynamic open-source software initiatives are leading the advances in this field.

Among many, the following **barriers** have been identified as principal. Those should be used as the main driving forces to orient research during the period 2025-2027 in Europe:

- There is a lack of coverage withing the range of computing devices in the continuum. Non-Linux equipment, IoT-resource-constrained elements, FPGAs and others are often disregarded, thus the continuum does not stretch well to those kinds of devices.

- There is no proper open source, end-to-end service pipelines that abstract the complexity and heterogeneity of the underlying equipment.
- The management of the lowest layers of the continuum is not solved at all, with connectivity and interoperability issues even at the level of research actions.
- Privacy and security are not tackled homogeneously across the entire continuum, and all stages of applications' lifecycle and may require a holistic approach from different clusters (i.e. CL4 and CL3).
- The lack of formal structure of the continuum hinders potential of organization and standardization.

Research challenges

Drawing from the previous, UPV and SRIPAS believe that the following challenges (of in short-term and mid-term priorities) should be explicitly tackled in the next Cluster 4 work program:

Short-term priorities:

- Combined hierarchical-P2P continuum content caching and distribution.
- Advanced (frugal, explainable, trustworthy) AI mechanisms for improving continuum performance: robustness, autonomy, traceability, governance
- Underlying network automation, IP-abstraction and connectivity by service names (eBPF, kernel technologies…).
- Achievement of a true AI-based cognitive mesh with ambient intelligence.
- Code-aid AI-powered support tools for continuum workloads generation, including CI/CD.
- Miniaturization of workloads' containers and packages.
- Explainable offloading and orchestration, involving context-aware self-configuration of workloads (from day-0 to day-2).
- Dynamic adjustment of performance and resource utilization based on real-time workload and environmental conditions, maximizing efficiency and minimizing energy consumption.
- Further development of mechanisms for autonomous detection, diagnosis, and correction of hardware and software failures, ensuring uninterrupted service and enhanced system resilience.
- Automatic semantic annotation of data, and autonomous metatada and schema discovery.
- Business models for effective federation of resource sharing in multi-stakeholder scenarios.
- Collaborative security threat modelling, identification, prediction, prevention and mitigation in a continuum formed by a computing swarm.
- Haptic codecs shared through the continuum for zero-latency, tactile services distribution.
- End-to-end benchmarking tools and methodologies for continuum environments in: self-capabilities, security & privacy, efficiency, usability, interoperability and sustainability.

Mid-term priorities:

- One-size-fits-all abstracted compilers for automatic, ML-based code generation towards global compatibility (all compilation targets in the continuum).
- Lightweight DLT to govern all continuum transactions (sovereign continuum).
- Alternative CPU architectures completely embedded in the continuum, prominently RISC-V.
- Auto-adapting communications: smooth L2 up to L7 automatic protocol swapping.
- Reconfiguration of the continuum in runtime in milliseconds time.
- Focus on hardware: electronic printing materials, energy harvesting, etc.
- Metaverse of the continuum, including VR simulation and IoT-powered ubiquitous data. Self-organizing networks with automatic formation, maintenance, and adaptation to ensure optimal data flow, load balancing, and resource allocation in complex, distributed computing environments.

In addition, we believe that interdisciplinarity of the research in the continuum must be reinforced during the coming years, incorporating investigation on (prominently) trustworthy AI, energy efficiency (e.g., batteries), environment (e.g., shared continuum footprint) and Social Science and Humanities (e.g., human-robot interaction), among others.