ACES – Autopoietic Cognitive Edge-cloud Services

Luca Alessandro Remotti[1], Chiara Senfett[1] and Maria Patrizia Meloni[1]

[1] *Datapower Consulting, Via Baldassarre Castiglione, 55, 00142, Roma (RM), Italy*

Abstract

The increasing need for cloud services at the edge (edge–services) is caused by the rapidly growing quantity and capabilities of connected and interacting edge devices exchanging vast amounts of data. This poses different challenges to cloud computing architectures at the edge, such as i) ability to provide end-to-end transaction resiliency of applications broken down in distributions of microservices; ii) creating reliability and stability of automation in cloud management under increasing complexity iii) secure and timely handling of the increasing and latency sensitive flow (east-west) of sensitive data and applications; iv)need for explainable AI and transparency of the increasing automation in edge-services platform by operators, software developers and end-users. ACES will solve these challenges by infused autopoiesis and cognition on different levels of cloud management to empower with AI different functionalities such as: workload placement, service and resource management, data and policy management. ACES key outcomes will be: i) autopoiesis cognitive cloud-edge framework; ii) awareness tools, AI/ML agents for workload placement, service and resource management, data and policy management, telemetry and monitoring; iii) agents safeguarding stability in situations of extreme load and complexity; iv) swarm technology-based methodology and implementation for orchestration of resources in the edge; v) edge-wide workload placement and optimization service; vi) an app store for classification, storage, sharing and rating of AI models used in ACES. ACES will be demonstrated and validated in 3 scenarios demanding for support of highly decentralised computing, ability to take autonomic decisions, reducing costs of cloud-edge management and increasing their efficiency, thus reducing impact on environment. To foster the uptake of ACES outcomes beyond its lifespan, different activities are foreseen to drive adoption to a wider network of stakeholders in key sectors.

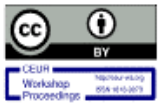
Keywords

Edge to cloud continuum, edge computing, autopoiesis, cognition

1. Vision and objectives

The ongoing escalation in data volume within organizational, service, and industrial processes in our daily living and working spaces can greatly benefit from a robust edge computing framework. This framework is tailored to process Big Data and AI locally, supporting a distributed architecture that fosters the interoperability of increasingly autonomous and intelligent applications and platforms. As a result, the initial cloud-to-edge (North-South) approach that marked the beginning of the 4th Industrial Revolution is evolving into an edge to-edge-to-cloud (East-West) model, priori-zing an 'edge-first' strategy. The development and management of this infrastructure are progressing in parallel with the applications and platforms it supports, emphasizing autonomous, 'edge-first' east-west interoperability, and distributed intelligent behaviors, sometimes showing emergent characteriscs. ACES is at the forefront, developing the inaugural platform for managing a powerful edge as a service (PEaaS) infrastructure. It is a high-performance, low-latency, and highly available platform that utilizes cognitive technologies to efficiently orchestrate and manage clusters of hyperconverged servers within a network of compact
EdgeMicroDataCenters (EMDC). ACES is characterized by its self-organizing, self-configuring, self-healing, and self-tuning intelligence, leveraging the full spectrum of Artificial Intelligence, Machine Learning, and Swarm Intelligence to maintain a self-reliant edge infrastructure. Its design is future-ready, capable of managing a fully composable data center equipped with the latest hardware, including PCIe CXL technology. The ACES autopoiesis framework represents a forward-thinking approach, offering the first generation of its platform capable of executing processing, storage, and networking services at the edge reliably and adaptively. This ensures ongoing self-optimization, minimal overprovisioning or underutilization, and reduced data transfer needs, making ACES a key player in achieving cost-effective, high-performance edge infrastructures.

Proceedings EGOV-CeDEM-ePart conference, September 1-5, 2024, Ghent University and KU Leuven, Ghent/Leuven, Belgium
luca.remotti@data-power.net (L.Remotti); chiara.senfett@data-power.net (C. Senfett), patrizia.meloni@data-power.net (P. Meloni)



ACES incorporates advanced containerization tools like Kubernetes, enhanced with cutting edge swarm technology for emergent intelligence, AI/ML. Unlike typical swarms, ACES has engineered a diverse swarm by amalgamating a large number of entities from uniquely developed species, initially cloning and merging traits from various well-known swarm species. The additional AI/ML is designed to provide the swarm with inputs, fine-tune the swarm agents, and introduce explainability into the swarm's behavior. This design aims to surpass the limitations of current orchestration mechanisms and usher in the future of data centers with composable hardware. It also seeks to enhance the efficiency of the standard Kubernetes orchestrator under rapidly changing workloads through the use of specific AI/ML and swarm technologies. While AI/ML has been somewhat slow in adapting to the quick shifts in workloads, and swarm technologies, though faster, have yet to offer sufficient adaptability or explainability, the use of these technologies has shown improved efficiency over standard Kubernetes orchestrators. The cutting-edge implementation of a dual high-speed fabric (Ethernet and PCIe) in EMDCs allows for dynamic allocation of CPU cores, accelerators, and storage, enabling unprecedented flexibility in hardware configuration for incoming workloads. Although not all hardware is currently CXL compatible, ACES is pioneering the orchestration for a fully composable data center, especially beneficial at the edge where resources are limited, enhancing utilization levels, energy efficiency, and ultimately reducing costs of edge infrastructure.

2. Key results

ACES undertakes research and technological innovation to respond to the increasing need of edge-cloud computing and data management and the demand of edge services. ACES edge cloud data and application services have the potential to enable a new infrastructure model, capable of guaranteeing end-to-end transaction resilience. The ACES solution provides autonomy and self-regulating mechanisms that provide systems stability, locally and edge wide. The requirements include the need for a horizontal flow of data and applications between sites as well as tackling issues of bandwidth, energy efficiency, security, and privacy. Furthermore, the autonomous operations on the platform need to be clearly explainable to operators, application developers and end-users and low-overhead is required in terms of costs, latency, energy, labour. ACES will provide an edge-services cloud with hierarchical intelligence, specifically autopoiesis and cognitive behaviours to manage and automate the platform. The project will carry out the following key outcomes: autopoietic edge-cloud data and application service platform; management agents and tools for awareness; AI and M-L enabled tools to handle workload, service and resource management, data and policy management, telemetry and monitoring; autopoietic agents for service and throughput stability in challenging service scenarios in terms of workload placement, service and resource management, data and policy handling; swarm technology-based solutions for orchestration of resources in the edge; edge-wide workload placement and optimization service; app store for ACES AI models classification, storage, sharing and rating.

Some examples of applications to be built are as follows:

• **The energy marketplace case study** in Greece demonstrates how distributed edge services can autonomously match energy supply and demand across regions, promoting renewable energy use and optimizing resource distribution.

• **Distributed energy grid process management**: utilizing an edge mesh for the Greek energy grid decentralizes management, enhancing the use of local energy resources and adapting to consumption needs, shifting from centralized to a resilient, adaptive infrastructure with a user interface that aids operators in decision-making and intervention.

• **An IoT based asset monitoring and management**: the innovation aims to show that integrating Advanced Metering Infrastructure data, grid-edge sensors, and GIS systems can enhance outage detection, improve prediction accuracy, and support reliable investment planning, including deferral, by analyzing diverse IoT and operational data for asset life assessment.

3. Key impacts

The aim of ACES is to develop a distributed, opportunistic, collaborative, heterogeneous, self-managed, self-organizing edge services environment, primarily edge-to-edge and secondly on the edge-to-cloud continuum.

The expected impacts of this implementation are:

• Improved placement of Europe in the delivery of secured edge-cloud service platforms in the global scenario;
• A reinforced capability of Europe to have available technical, computational and data transmission means to manage urgent societal challenges;
• Availability of more effective technologies and tools to manage distributed cloud systems at the edge.

More specific impacts of ACES concern:

• The energy sector, facilitating the transition towards a system capable of optimising the relationship between supply and demand and the integration of sustainable energy sources;
• The more general impact on the European Green Deal, driving the concept of smart infrastructure and decentralised energy production;
• Impact on sustainable development goals;