

D4.9 COGNIT Serverless Platform - Software Source - d

Version 1.0

30 April 2025

Abstract

COGNIT is an open-source serverless framework for the cognitive cloud-edge continuum that enables the seamless, transparent, and trustworthy integration of data processing resources from providers and on-premises data centers in the cloud-edge continuum. Automatic and intelligent adaptation to where and how data is being processed according to application requirements can be achieved by enabling an AI-Enabled orchestrator. Further, the orchestrator is able to optimise the placement, acts when changing in application demands and behaviour, and the operation of the infrastructure in terms of the main environmental sustainability metrics. This document provides a catalogue of those open source software components developed in WP4 “AI-Enabled Distributed Serverless Platform and Workload Orchestration” during the Fourth Research & Innovation Cycle (M22-M27) as part of the implementation of several of the main components of the COGNIT Framework (i.e. Cloud-Edge Manager and AI-Enabled Orchestrator).

Deliverable Metadata

Project Title:	A Cognitive Serverless Framework for the Cloud-Edge Continuum
Project Acronym:	SovereignEdge.Cognit
Call:	HORIZON-CL4-2022-DATA-01-02
Grant Agreement:	101092711
WP number and Title:	WP4. AI-Enabled Distributed Serverless Platform and Workload Orchestration
Nature:	R: Report
Dissemination Level:	PU: Public
Version:	1.0
Contractual Date of Delivery:	31/03/2025
Actual Date of Delivery:	30/04/2025
Lead Author:	Monowar Bhuyan (UMU) & Paul Townsend (UMU)
Authors:	Yashwant Singh Patel (UMU), Adil Bin Bhutto (UMU), Rohail Gulbaz (UMU), Malik Bouhou (CETIC), Javad Forough (UMU), Aritz Brosa (Ikerlan), Idoia de la Iglesia (Ikerlan), Sébastien Dupont (CETIC), Fátima Fernández (Ikerlan), Joan Iglesias (ACISA), Tomasz Korniluk (Phoenix), Johan Kristiansson (RISE), Antonio Lalaguna (ACISA), Marco Mancini (OpenNebula), Philippe Massonet (CETIC), Nikolaos Matskanis (CETIC), Aitor Garciandia (Ikerlan), Mikel Irazola (Ikerlan), Álvaro Puente (Ikerlan), Juan José Ruiz (ACISA), Kaja Swat (Phoenix), Thomas Ohlson Timoudas (RISE), Iván Valdés (Ikerlan), Victor Palma (OpenNebula), Michal Opala (OpenNebula), Jakub Walczak (OpenNebula), Carlos Moral (OpenNebula)
Status:	Submitted

Document History

Version	Issue Date	Status ¹	Content and changes
0.1	20/04/2025	Draft	Initial Draft
0.2	23/04/2025	Peer-Reviewed	Reviewed Draft
1.0	30/04/2025	Submitted	Final Version

Peer Review History

Version	Peer Review Date	Reviewed By
0.2	23/04/2025	Antonio Álvarez (OpenNebula)
0.2	23/04/2025	Aritz Brosa (Ikerlan)

Summary of Changes from Previous Versions

First Version of Deliverable D4.9

¹ A deliverable can be in one of these stages: Draft, Peer-Reviewed, Submitted, and Approved.

Executive Summary

This is the fourth “COGNIT Serverless Platform - Software Source” report that has been produced in WP4 “AI-Enabled Distributed Serverless Platform and Workload Orchestration”. It provides a short description, licence, version, code repository and user guide, as well as the design, testing, and verification references of each of the software requirements that have had active development tasks during the Fourth Research & Innovation Cycle (M22-M27) in connection with these main components of the COGNIT Framework:

Cloud-Edge Manager

- **SR4.6** Plan Executor:

Plan Executor is responsible for the execution of plans produced by the AI-Enabled Orchestrator related to the placement and migration of Serverless Runtimes within an Edge Cluster.

AI-Enabled Orchestrator

- **SR5.1** Building Learning Models:

Provide AI/ML models based on collected metrics from the Cloud-Edge Manager monitoring service related to Edge Clusters, Serverless Runtimes, and infrastructure usage.

- **SR5.2** Smart management of Cloud-Edge resources:

AI-Enabled Orchestrator is responsible for managing and optimizing the lifecycle of Edge Clusters and serverless runtimes within Edge Clusters according to the application requirements, infrastructure and virtual resource usage, and energy-aware policies.

This deliverable has been released at the end of the Fourth Research & Innovation Cycle (M27), and will be updated with a standalone version in M33.

Table of Contents

Abbreviations and Acronyms	5
1. Cloud-Edge Manager	6
2. AI-Enabled Orchestrator	7

Abbreviations and Acronyms

AI	Artificial Intelligence
API	Application Programming Interface
CPU	Central Processing Unit
FaaS	Function as a Service
MOGA	Multi-Objective Genetic Algorithm
ML	Machine Learning
NSGA-II	Non-dominated Sorting Genetic Algorithm II
REST	Representational State Transfer
SPEA2	Strength Pareto Evolutionary Algorithm 2
SR	Serverless Runtimes
VM	Virtual Machine

1. Cloud-Edge Manager

[SR4.6] Plan Executor	
Description	Plan Executor component is responsible for the execution of plans produced by the AI-Enabled Orchestrator related to the placement and migration of Serverless Runtimes within an Edge Cluster. Plans are implemented by performing actions on the Serverless Runtimes such as deploy, migrate, poweroff, etc.
Licence	Apache 2.0
Version	OpenNebula 7.0
Design	D4.4 → SR4.6 Plan Executor
Code	OpenNebula Public Repository
User Guide	OpenNebula Documentation
Testing	D5.5 → 9.4 Cloud-Edge Manager
Verification	D5.5 → 9.4 Cloud-Edge Manager

2. AI-Enabled Orchestrator

[SR5.1] Building Learning Models

Description The AI-Enabled Orchestrator utilises machine learning models developed (e.g., classical K-means, Deep K-means) for unsupervised workload classification. The classification results are combined with scheduling algorithms to optimise the serverless runtimes across cloud-edge resources, enhancing resource interference, carbon emission, and monetary cost. The orchestrator leverages the auto-adaptive ML retraining method to mitigate the impact of data drift to refine its learning models adaptively.

Licence Apache 2.0

Version 3.0

Design D4.4 → [SR5.1] Building Learning Models

Code [Public Repository](#)

User Guide [Repository README](#)

Testing D5.5 → 9.5 AI-Enabled Orchestrator

Verification D5.5 → 9.5 AI-Enabled Orchestrator

[SR5.2] Smart management of Cloud-Edge resources

Description The AI-enabled orchestrator is responsible for managing and optimising the lifecycle of Edge Clusters and serverless runtimes within Edge Clusters according to the application requirements, infrastructure and virtual resource usage, and energy-aware policies. The orchestrator leverages the optimal SR estimator, initial placement of SRs, and multi-objective optimization algorithms (e.g., NSGA-II, MOGA, and SPEA-2) to balance resource interference, carbon emission, and monetary cost across Cloud-Edge clusters.

Licence Apache 2.0

Version 2.0

Design D4.4 → [SR5.2] Smart Management of Cloud-Edge Resources

Code [Public Repository](#)

User Guide	Repository README
Testing	D5.5 → 9.5 AI-Enabled Orchestrator
Verification	D5.5 → 9.5 AI-Enabled Orchestrator
