Grant Agreement No.: 101070473

Call: HORIZON-CL4-2021-DATA-01

Topic: HORIZON-CL4-2021-DATA-01-05

Type of action: HORIZON-RIA

D7.1 MARKET ANALYSIS

Work package	WP 7
Task	T7.1
Due date	November 2023
Submission date	27.11.2023
Deliverable lead	BOR
Version	1.0
Authors	Simon Hinterholzer Isabel Merz Ralph Hintemann Severin Beucker
Reviewers	Costanza Pestarino Giovanni Rimassa

Abstract	This document puts the focus on analysing technologies and market offerings related to FLUIDOS and its associated use cases. To this end, the document includes a matrix for analysing related market offerings from various perspectives and analyses the results from some interviews with relevant actors to clarify/detail some of the aspects. The document also derives some first steps for a long term FLUIDOS exploitation.
Keywords	Computing continuum; liquid computing; FLUIDOS architecture

Funded by Horizon Europe Framework Programme of the European Union

Document Revision History

Version	Date	Description of change	List of contributor(s)
V0.1	28/09/2023	1st version of the report	BOR
V0.2	15/11/2023	Version prepared for review	BOR
V1.0	27/11/2023	Document for submission/publication	BOR/DSME/MAR

Disclaimer

The information, documentation, and figures available in this deliverable are written by the "Flexible, scaLable and secUre decentralIzeD Operating System" (FLUIDOS) project's consortium under EC grant agreement 101070473 and do not necessarily reflect the views of the European Commission.

The European Commission is not liable for any use that may be made of the information contained herein.

Copyright notice

Project co-funded by the European Commission in the Horizon Europe Programme		
Nature of the deliverable:	R	
Dissemination Level		
PU	Public, fully open, e.g. web	☒
SEN	Sensitive, limited under the conditions of the Grant Agreement	
Classified R-UE/ EU-R	EU RESTRICTED under the Commission Decision No2015/ 444	
Classified C-UE/ EU-C	EU CONFIDENTIAL under the Commission Decision No2015/ 444	
Classified S-UE/ EU-S	EU SECRET under the Commission Decision No2015/ 444	

** R: Document, report (excluding the periodic and final reports)*

DEM: Demonstrator, pilot, prototype, plan designs

DEC: Websites, patents filing, press & media actions, videos, etc.

DATA: Data sets, microdata, etc

DMP: Data management plan

ETHICS: Deliverables related to ethics issues.

SECURITY: Deliverables related to security issues

OTHER: Software, technical diagram, algorithms, models, etc.

Funded by Horizon Europe Framework Programme of the European Union

TABLE OF CONTENTS

Funded by Horizon Europe Framework Programme of the European Union

ABBREVIATIONS

AI	Artificial Intelligence
API	Application Programming Interface
AR	Augmented Reality
ARPANET	Advanced Research Projects Agency Network
ASF	Apache Software Foundation
AWS	Amazon Web Services
CAGR	Compound Annual Growth Rate
CDN	Content Delivery Network
CPU	Central Processing Unit
CSP	Cloud Service Provider
DAI	Distributed Artificial Intelligence
EC2	Amazon Elastic Compute Cloud
GB	Gigabyte
GCP	Google Cloud Platform
GNU GPL	GNU General Public License
GPU	Graphics Processing Unit
HPC	High Performance Computing
IaaS	Infrastructure as a Service
ICCC	International Conference on Computer Communications
IOT	Internet of Things
IT	Information Technology
MT/s	Megatransfers per second
MW	Megawatt
PaaS	Platform as a Service

Funded by Horizon Europe
Framework Programme of the European Union

RAM	Random Access Memory
SaaS	Software as a Service
SME	Small and Medium-sized Enterprise
TPU	Tensor Processing Unit
TRL	Technologyl Readiness level
TWh	Terawatt hours
vCPU	Virtual Central Processing Unit
VM	Virtual Machine
VRAM	Video Random Access Memory
VR	Virtual Reality

Funded by Horizon Europe Framework Programme of the European Union

1 INTRODUCTION

1.1 A BRIEF LOOK AT THE HISTORY OF COMPUTING AND THE VISION OF FLUIDOS

In the early years of digital computing during the initial decades of the 20th century, the landscape was characterised by a small number of computing devices, which required vast amounts of space and energy in relation to their constrained computing capacity. However, the introduction of silicon-based transistor technology from the 1950s and 1960s facilitated a substantial surge both in the proliferation of computing devices and their performance capabilities, a trajectory that continues to evolve.

Already early in the history of computing, there emerged an idea that digital resources did not need to be limited solely to the confines of a physical computing device. The inherent nature of data and information, unlike material things, allowed for seamless transportability. For example, more than 50 years ago, at the International Conference on Computer Communications (ICCC) 1972, an early version of the Internet (ARPANET) showcased how the capacities/applications of distant computers could be used via a terminal (Robert Kahn, 1972).

The development of wide area networks and the Internet has firmly established the client-server infrastructure for both technical prowess and economic reasons. Essentially, similar to the ARPANET, less complicated and cost-effective clients or personal computers, formerly known as terminals, leveraged the resources of high-powered computers or servers over the network for data and computing needs.

Consequently, this evolution has paved the way for the development of data centres, which provide protected and stable environments to house many servers. As a result, the expansion of the Internet and the escalation in network speeds have led to create entirely new applications (social media, video streaming, etc.).

In addition to the above, there has been a surge in the growth of cloud technology, marked by a continual increase in the construction and operations of hyperscale data centres. However, the IT hardware that powers data centres depends upon a wide range of materials, triggering social and environmental concerns during their production. In addition, the energy consumption of data centres is increasing, with Europe set to consume nearly 80 TWh in 2018 (Hintemann, Hinterholzer, Montevecchi, & Stickler, 2020). Although there is still much uncertainty in the scientific community about the energy and material consumption of data centres and cloud computing, numerous studies, such as the one conducted by the International Energy Agency, highlight a discernible upward trend in energy demand (Vida Rozite, Emi Bertoli, & Brendan Reidenbach, 2023).

Cloud technology, which is based on virtualization, facilitates greater utilisation of IT hardware, thereby enabling more environmentally friendly operations compared to many traditional, partially application-specific data centres and server rooms. With the virtualization of hardware (vCPU, vRAM etc.), virtual machines and containers were also established, which organize applications together with the application environment in specific packages that can be operated in a highly distributed manner on this virtual hardware and can be dynamically scaled and moved. For economic reasons, this has led to the consolidation of computing capacity in hyperscale cloud computing centres in recent years.

Software designed specifically for such virtualized and distributed computing environments (cloud native), follows a number of paradigms to take full advantage of this infrastructure. Instead of a central

Funded by Horizon Europe Framework Programme of the European Union

monolithic application, it deploys a variety of software components that work together through well-defined interfaces. Should a software component need more compute, RAM, or storage capacity due to an expanding user base, it can automatically be provisioned with more vCPUs, VRAM or storage in a cloud environment. To ensure that the individual software components work together reliably, well-defined application programming interfaces (APIs) are required. At the same time, redundancy requirements have changed, thus cloud native applications can be moved virtually between data centres during operation to reduce the need for single data centre reliability and can reduce traditional redundancy requirements (such as n+1, 2n or 2n+1) for power and cooling. If one data centre fails, a service is simply moved to another data centre of the same Cloud Service Provider on the software side.

The advancements in virtualizing and automating cloud resources, coupled with the construction and utilization of cloud software leveraging these virtual resources, have enabled applications to operate across an extensively distributed infrastructure. In contrast to the development of hyperscale clouds, at the other end of the Internet, close to the end device or user (the so-called edge), current trends are leading to an enormous increase in the number of growingly powerful connected devices. These are not just the notebooks, tablets, and smartphones we all know and use, but also robots, embedded systems, billions of sensors and much more - the aptly named Internet of Things (IoT). With multi-core CPUs and GPUs, many of these devices have far more computing power than they need for their day-to-day tasks - more than 95% of the time, many of the physical resources are typically not being utilized. (reference)

This is the starting point for the vision of FLUIDOS, an edge cloud operating system, capable of providing such highly distributed capacity in a continuum comparable to central clouds. Naturally, the system can take into account that the actual core function of the end device should never be impaired by external processes, which should only run when there is sufficient surplus capacity. With end-to-end networking, applications running in this continuum can be migrated dynamically. Finally, the continuum should not be confined to the highly distributed devices, but should potentially be able to dynamically move workloads to traditional clouds according to multi-criteria (e.g., economic and environmental) objectives.

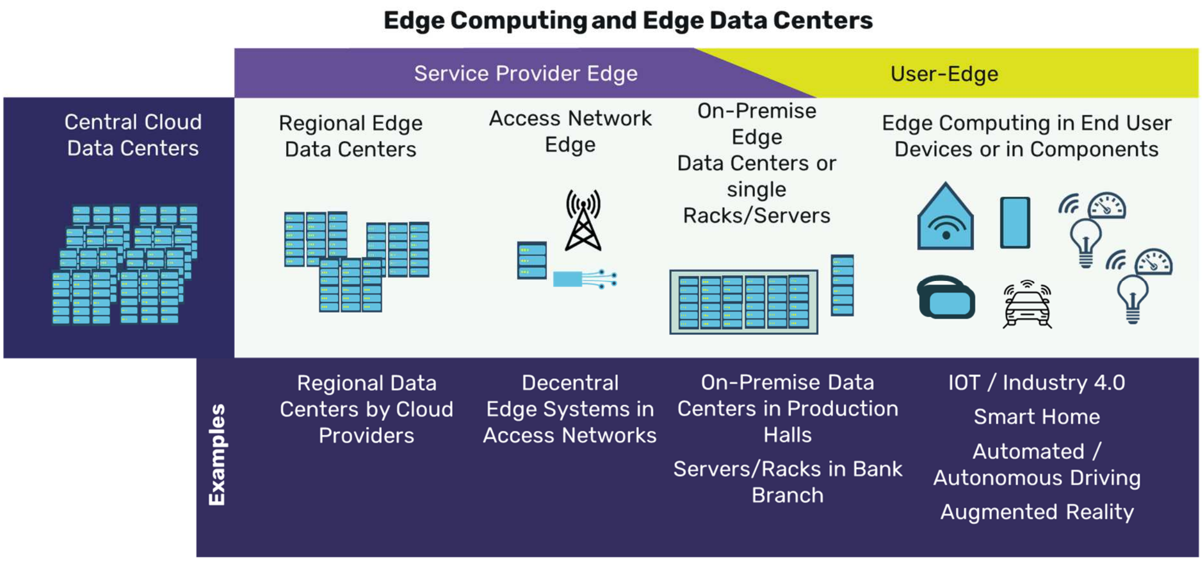

Figure 1: Different types of edge computing, based on (Hintemann, Hinterholzer, & Grothey, 2022)

Software systems such as Kubernetes are designed to manage and scale container-based applications across disparate environments, including the edge. Software-based resilience, combined with critical

Funded by Horizon Europe Framework Programme of the European Union

infrastructure, ensures a seamless transition in the event of unexpected occurrences. It is all about building and designing cloud environments in a way that downtime is not only minimized, but in most cases removed from the equation altogether (Marshall, 2021).

1.3 Market projections for edge computing and fields of application

In the recent past, strong growth has been predicted for edge computing, mainly based on the forecasts of renowned analysts and market research institutes such as Gartner, IDC or McKinsey (Bittman, 2017; JM Chabas, Chandra Gnanasambandam, Sanchi Gupte, & Mitra Mahdavian, 2018; Shirer, 2021; van der Meulen, 2019). In this context, Gartner (van der Meulen, 2019) has notably revised its forecasts for the edge's share of enterprise data processing several times to the effect that growth is more likely to occur later. These developments can be tracked back in the Internet Archive's Wayback Machine ('Internet Archive', 2023).

How quickly and to what extent edge devices and infrastructures will expand is still unclear from today's perspective. On the one hand, the development of key technologies that are considered to be drivers of edge computing, such as autonomous driving, IoT, AI, augmented and virtual reality, is not yet precisely foreseeable. Additionally, it is often questionable how edge structures will be set up in practice. How much distributed edge infrastructure will be needed on the roads for autonomous driving and how much computing power can be done directly in the vehicle (Bonomi, Poledna, & Steiner, 2017)? This kind of question can be asked for various applications. The forecasts for the edge computing market often exhibit significant disparities and variations (Global Market Insights, 2020; Marshall, 2021; Shirer, 2021).

Nevertheless, it has become further and further apparent in many places that edge computing is significantly gaining in importance. Complex end-user devices/systems are getting more and more computing power to process increasing amounts of locally generated data. Examples can be found in areas like vehicles with assistance functions, intelligent building and home control, but also in professional areas such as production facilities (smart factory), medical robotics and real-time optimization in logistics (Abbas, Zhang, Taherkordi, & Skeie, 2017; Cao, Liu, Meng, & Sun, 2020; Denby & Lucia, 2020; Khan et al., 2020; Shi & Dustdar, 2016). Collaborations between network operators and cloud service providers, who are currently installing edge infrastructures at distributed network points, are also increasingly observed (Edge Computing, 2020; van der Meulen, 2019; Vodafone, 2021, among others).

Global Market Insights forecasted a growth rate for the edge data centre market of over 15% CAGR until 2032. (Global Market Insights, 2022) Other sources estimated the growth of the market to be much higher: Grand View Research forecasted a development of the worldwide edge computing industry even at a CAGR of 38.4% in the timespan from 2021 to 2028 (Grand View Research, 2022). Besides, significant increases in spending within the edge computing sector are anticipated. For instance, projections suggested a 13% annual increase to $317 billion by 2026 (IDC Corporate, 2023)).

In "The 2021 State of the edge report" from the Linux Foundation, cumulative capital expenditures of up to $800 billion were predicted for edge computing in the time frame between 2019 and 2028. The report highlighted an increasing demand for edge infrastructure since the pandemic (The Linux Foundation, 2021).

Omdia further assessed that revenues for industrial edge networking components (like connectors, protocol converters etc.) were likely to grow by 10.3% CAGR until 2025 (Omdia, 2021).

Funded by Horizon Europe Framework Programme of the European Union

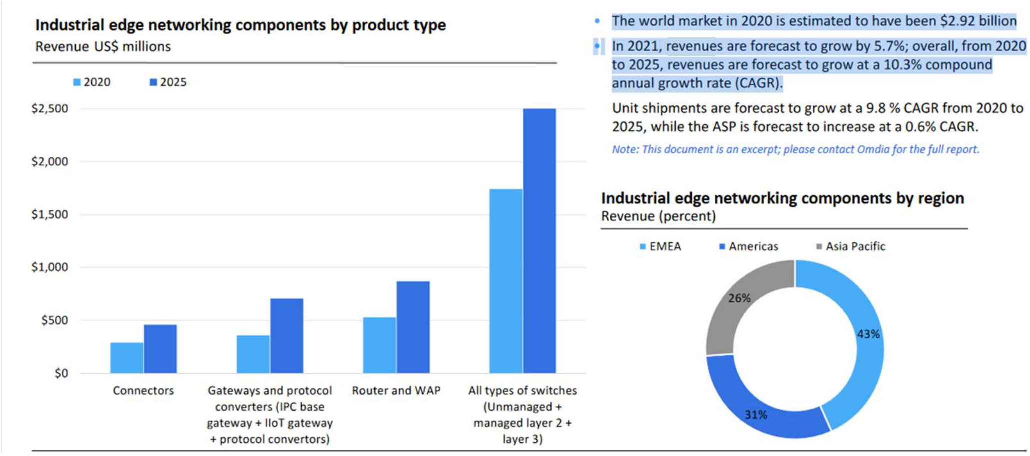

Figure 2, Revenues for industrial edge networking components (Omdia, 2021)

Manufacturing applications in the European edge data centre market are expected to grow at a CAGR of 20 percent by 2026 (DataCenter Insider, 2021). Regarding the installed capacity in terms of MW, in "The State of the Edge report" for the Linux foundation it is stated that the global IT power footprint in relation to edge infrastructure is supposed to increase by more than 40 times until 2028, with an annual growth rate of 40 % (Marshall, 2022).

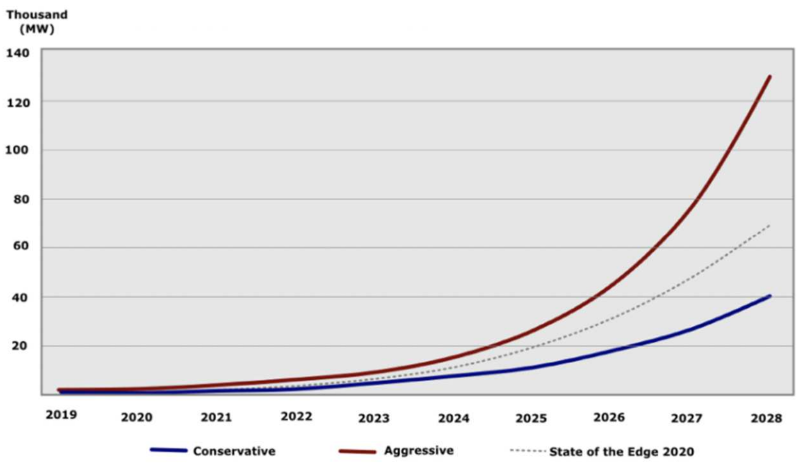

Figure 3: Growth of edge computing capacity based on "State of the edge report 2022" by Linux foundation (Marshall, 2022)

Gartner´s forecast indicated that by 2025, approximately 75% of enterprise data would not be processed within centralized data centres/cloud but in edge solutions (Gartner, 2018). At the same time, there are other evaluations, such as the one from Ars Technica, which considered the development of distributed computing (here especially volunteer distributed computing) at its ending. Rising power costs, the shift from desktop computers to laptops with lower CPUs and a lack of public interest in volunteer computing projects are perceived as obstacles for the future development of edge deployments (Ars Technica, 2023).

Funded by Horizon Europe Framework Programme of the European Union

2 AIM OF THIS REPORT AND METHODS USED

This report focuses on the analysis of technologies and market offerings related to FLUIDOS and its associated use cases. The report will create a matrix for evaluating comparable market offerings from different perspectives. Additionally, the report will employ technical radars as visual tools to differentiate FLUIDOS´ technical approach from both existing solutions and alternative technical approaches. This will entail comprehensive research into technologies and current offerings. Moreover, it will involve conducting targeted interviews with key stakeholders to delve into specific aspects and conduct a more in-depth analysis of potential markets for FLUIDOS. The final chapter builds on this and attempts to provide information for further exploitation and dissemination of the results and findings of the FLUIDOS project. The methods used to prepare this report include:

<u>Desk research:</u>

Extensive online research was conducted for similar market offerings and technologies. Scientific literature (various open access journals), targeted trade journals, websites, and blogs were searched. In particular, Google, Google Scholar and ResearchGate searches were used. Some main keywords used in different combinations for the search include:

> [distributed, edge, fog, meta, operating, system, fluid, network, continuum, grid, resource, sharing, cloud, computing, Kubernetes]

The results were selected and filtered based on the relevance and/or connection to FLUIDOS´ mission. A relatively generous approach was taken as the exact characteristics of FLUIDOS, as an edge-cloud continuum with a zero-trust paradigm, horizontal and vertical resource sharing, intent-based service and inclusion of almost all device types/operating systems, does not exist in this form. There was no restriction on the publication date of the sources. The classification information found was extracted directly from the source documents or, where necessary, a separate search was conducted.

<u>Interviews:</u>

Five interviews were conducted to better differentiate the uniqueness of FLUIDOS from other existing solutions as well as new approaches, and to better understand the role and added value of FLUIDOS in specific use cases or sectors.

Structured guiding questions were conducted with the use case owners within FLUIDOS as well as with an external expert (specialized in distributed computing in the healthcare sector) specific to the project.

With Topix, the perspective was explained from the point of view of an internet exchange involved in FLUIDOS and engaged in federated network and cloud services. Polito was also interviewed to better understand the vision and intent behind FLUIDOS.

The following guiding questions were asked of the participants. However, depending on the course of the interview and the objective, slight adaptations were made, and additional inquiries were spontaneously introduced:

- What is your view on the development of a cloud-edge continuum? What are your expectations and concerns in general?

Funded by Horizon Europe Framework Programme of the European Union

- As a research project, FLUIDOS is still at a very early stage (the goal is to reach TRL 5 by the end of the project). Can you imagine that the company you work for (alternatively your industry) will use FLUIDOS in the future?
- Who else do you think would use FLUIDOS to a) provide cloud services or b) use cloud services running on FLUIDOS (or both)?
  - For what purposes or applications?
- Would FLUIDOS replace or complement existing software/technology?
- What do you think are the advantages of FLUIDOS compared to the current solution your company/sector uses today?
- Which features of FLUIDOS would be essential for the use of FLUIDOS in your company? (If not mentioned – ask about zero trust, resilience through workload shifting in case of failure of individual nodes, workload shifting according to ecological criteria...)
- What could be possible disadvantages of using FLUIDOS in your company?
- Are you aware of other similar technical approaches that are trying to enter the market?
- What do you think is important for FLUIDOS to reach a wide user base (in your sector / in general)?

Funded by Horizon Europe Framework Programme of the European Union

3. Market segmentation and different user- and application types

In the following chapter, on the basis of the desk research, different types and segments of computing will be described. For the various market segments, existing offerings will be compared.

3.1 Market segmentation

The segments of the computing market, such as traditional data centres, cloud computing, and edge computing, can be categorized based on various factors like infrastructure, location, and service model. Below is a categorized breakdown of the segments:

Infrastructure:

Traditional data centres: These are typically large, centralized facilities that house and manage servers, storage, and networking equipment. They are usually owned and operated by businesses or data centre service providers.

Cloud computing: Cloud computing infrastructure is provided by cloud service providers (CSPs) and consists of a network of servers and data storage that is accessible over the internet. It includes Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) offerings.

Edge computing: Edge computing infrastructure is distributed and located closer to the data source or endpoint devices. It includes edge servers, gateways, and devices that process data locally before sending it to centralized data centres or the cloud.

Location:

Traditional data centres: These are typically centralized in specific geographical locations and may serve a wide area or even a global audience.

Cloud computing: Cloud data centres are often distributed across multiple regions and availability zones, strategically located to reduce latency and improve redundancy.

Edge Computing: Edge computing resources are distributed at the "edge" of the network, closer to where data is generated or consumed. This can include edge nodes in cities, factories, usual computers/smartphones or even on IoT devices.

Service Model:

Traditional data centres: These primarily offer infrastructure services that organizations manage themselves. Companies purchase and maintain their own hardware and software.

Cloud computing: Cloud services come in various service models:

- IaaS: Infrastructure as a Service provides virtualized computing resources, allowing users to manage and control the underlying infrastructure.
- PaaS: Platform as a Service offers a platform and environment for developers to build and deploy applications without worrying about infrastructure.

Funded by Horizon Europe Framework Programme of the European Union

- SaaS: Software as a Service delivers software applications over the internet, typically on a subscription basis.

Edge computing: Edge computing can support various service models, depending on the specific use case. It may involve edge devices processing data locally, edge servers running applications, or a combination of both.

Ownership and Management:

Traditional data centres: Typically owned and managed by individual organizations or data centre service providers.

Cloud computing: Managed by CSPs like Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP).

Edge computing: Can be owned and managed by a variety of entities, including organizations, telecom providers, and edge computing service providers.

These categories provide a framework for understanding the different segments of the computing market and how they serve different needs in the evolving landscape of IT infrastructure and services. Additionally, hybrid approaches combining elements of traditional data centres, cloud computing, and edge computing are becoming increasingly common as organizations seek to optimize their computing resources to meet their specific requirements.

A meta-operating system like FLUIDOS blurs the boundaries between these physically distinct categories. Since the physical properties behind the virtual resources are made transparent to the user or the application software, the best resources can dynamically be selected in an intent-driven manner. Given that the existing physical characteristics will always be reflected also in the virtualized resources, it is expected that performance-hungry applications (e.g. large in-memory databases) will continue to run on powerful servers as they do today in a data centre. Conversely, if the aim is to achieve the lowest feasible latencies for less intricate tasks, deploying a device at the edge may also prove advantageous. In addition to performance objectives, there are of course many other criteria, such as the availability of green power or redundancies for resilient operation, on which to base the selection of a compute cluster or node.

3.2 DIFFERENT USER- AND APPLICATION TYPES AND THEIR COMPUTING NEEDS

There are numerous application types in cloud computing, wherein the cloud hardware requirements and appropriate instance types vary depending on the application type. This section describes some common application types and their associated cloud hardware requirements and appropriate instance type, as well as the potential to provide such services or parts of it from an edge-cloud:

Web hosting and content delivery:

Cloud Hardware Requirements: Reliability and scalability are critical for delivering websites and content globally.

Appropriate instance types: General purpose VM instances with sufficient CPU, RAM and networking power to run web servers and CDNs.

Funded by Horizon Europe Framework Programme of the European Union

Edge cloud potential: Edge clouds can help bring content closer to end users to minimize latency and increase availability.

Data analytics and machine learning:

Cloud hardware requirements: High processing power and dedicated hardware accelerators (e.g. GPUs or TPUs) for compute-intensive tasks.

Appropriate instance types: GPU- or TPU-optimized instances for advanced data analytics and machine learning.

Edge cloud potential: Edge clouds can be useful for data-intensive applications to perform local pre-processing and analysis.

Internet of Things (IoT):

Cloud hardware requirements: Low latency, big data support and the ability to handle streaming data.

Suitable instance types: Edge computing devices or dedicated IoT instances with low-power CPUs and IoT-enabled communication interfaces.

Edge cloud potential: IoT applications are typical candidates for edge clouds to process data locally and reduce network load.

Collaboration and communication tools:

Cloud hardware requirements: High availability and low latency for real-time communications.

Suitable instance types: Instances with sufficient networking and processing power for real-time communication tools such as video conferencing and messaging applications.

Edge cloud potential: Edge clouds could help minimize latency for these applications.

Software as a Service (SaaS):

Cloud hardware requirements: Reliability and scalability to deliver SaaS applications to customers worldwide.

Suitable instance types: A wide range of instance types, depending on the requirements of the application being hosted.

Edge cloud potential: Certain SaaS services could benefit from edge clouds, especially where low latency or data protection requirements are important.

The suitability of an edge cloud depends on the specific requirements of the application. Applications that require real-time processing, low latency and the use of distributed endpoint resources could benefit from edge clouds. These include

- IoT applications: Networking and processing data from IoT devices is often done on edge clouds to minimize latency and reduce network traffic.
- Autonomous driving: Vehicles require fast local processing to make real-time decisions without accessing a remote cloud.
- Industrial automation: Edge computing is critical in factories and manufacturing plants to precisely control machines and process data locally.

Funded by Horizon Europe Framework Programme of the European Union

- Healthcare: Medical devices and patient monitoring can benefit from edge clouds to process data securely and in real-time.
- Smart cities: Sensors and cameras in urban infrastructure could use local edge clouds to make quick decisions for traffic control and security applications.

The choice between a central cloud and an edge cloud depends on the requirements and the specific use case. Often these two approaches complement each other as they have different strengths.

As you can see, depending on the (Edge-)devices and their individual characteristics (performance, chip architectures, RAM, storage, network connection, availability), a huge range of different applications can theoretically be run. A wide variety of devices and transparency of device capabilities would ensure the best possible quality of service with fully automated negotiation.

However, it is expected that some of the services will still be based on physical devices that are offering huge capacities (e.g. hundreds of GB of RAM) within one single physical Server to enable, for example, cross-analysis of huge amounts of data in a very short time. This would be extremely cumbersome with cumulative RAM connected via network to CPUs/GPUs, which is extremely slow compared to a memory bus with currently up to 8000 MT/s and latency as low as ~15-25 nanoseconds (Cutress, 2020). It follows that edge computing will probably never replace cloud computing, but it will effectively complement it and perhaps even replace parts of it in the long term. This is why FLUIDOS does not want to create a pure EDGE cloud, but it aims at integrating various devices at the edge as well as existing (more centralized) clouds into an edge-cloud continuum.

Funded by Horizon Europe
Framework Programme of the European Union

4 Market Environment and Technological Analysis

4.1 Market Environment: Related Technologies and Market Offerings

Based on the research described in Chapter 2, other existing technologies in development and market offerings that provide similar and relevant services for FLUIDOS were collected. The following table lists the aforementioned services by name, provides a brief description, whether they are open source, their development status and a source for further information.

Nr.	Name	Description	Open source	Status of development	Source/Link
1	Aarna Networks: Aarna Edge Services (AES)	Aarna Networks' solution provides zero-touch edge orchestration and management services. AES enables various tasks such as provisioning, configuring and managing multi-cloud connectivity for cloud edge users, colo users and SD-WAN service providers like Aryaka. Users can share primary or secondary/object storage, such as Seagate Lyve Cloud, across public clouds and edge applications without incurring egress costs, the company says. AES also allows Machine Learning applications to be used at the edge, reducing traffic to public clouds and eliminating the need for expensive on-prem hardware (Edge Industry Review, 2023).	open	Test phase/already in use	https://www.aarnanetworks.com/ https://www.edgeir.com/aarna-networks-launches-aes-a-zero-touch-orchestration-as-a-service-for-edge-computing-20230226
2	Accordion: Adaptive edge/cloud compute and network continuum over a heterogeneous sparse edge infrastructure to support nextgen applications	ACCORDION establishes an opportunistic approach in bringing together edge resource/infrastructures (public clouds, on-premise infrastructures, telco resources, even end-devices) in pools defined in terms of latency, that can support NextGen application requirements. To mitigate the expectation that these pools will be "sparse", providing low availability guarantees, ACCORDION will intelligently orchestrate the compute & network continuum formed between edge and public clouds, using the latter as a capacitor (ACCORDION project, 2022).		Research project	https://cordis.europa.eu/project/id/871793/de
3	aerOS	aerOS is a European research project (Horizon Europe CL4-2021-DATA-01-05) running for 3 years which aims at transparently utilising the resources on the edge-to-cloud computing continuum for enabling applications in an effective manner while incorporating multiple services. The overarching goal of aerOS is to design and build a virtualized, platform-agnostic meta operating system for the IoT edge-cloud continuum (Aeros, 2023).	open	Research phase (parallel research project to FLUIDOS)	aeros-project.eu

Funded by Horizon Europe Framework Programme of the European Union

4	AI EdgeLabs Sensor	AI EdgeLabs is a powerful and autonomous cybersecurity AI platform that helps security teams respond immediately to ongoing attacks and protect Edge/IoT infrastructures against malware, DDoS, botnets, and other threats. The EdgeLabs Sensor is a proprietary network telemetry and monitoring agent that continuously analyses and models network behaviour of the particular Edge node. It holds an AI-powered model stack that continuously checks inbound and outbound connectivity through pre-trained AI models. With EdgeLabs Sensor, infrastructure teams can locate hidden network threats and zero-day attacks that might cause business disruption. The sensor models normal and abnormal connectivity states and immediately reports them to Operation teams with cause and recommends remediation actions for the incidents (Edge labs, 2023).		Already in use	https://edgelabs.ai/platform/
5	Amazon EC2	Amazon Elastic Compute Cloud (Amazon EC2) offers the broadest and deepest compute platform, with over 600 instances and choice of the latest processor, storage, networking, operating system, and purchase model to help you best match the needs of your workload. We are the first major cloud provider that supports Intel, AMD, and Arm processors, the only cloud with on-demand EC2 Mac instances, and the only cloud with 400 Gbps Ethernet networking. We offer the best price performance for machine learning training, as well as the lowest cost per inference instances in the cloud. More SAP, high performance computing (HPC), ML, and Windows workloads run on AWS than any other cloud (AWS, 2023).		Widely used	https://aws.amazon.com/ec2/?nc1=h_ls
6	Amazon Luna	Edge Gaming: Amazon Luna is a cloud gaming service that lets you play games from the cloud on devices you already own. With cloud gaming, you don't need gaming consoles or gaming PCs. Games are streamed directly from the cloud to your screen, so you can play anywhere you have high-speed Internet access. Playing games on Luna with the Amazon Luna controller provides a low-latency gaming experience. A touch controller is also available on supported Fire tablets and mobile devices when a Luna controller is not connected to the device (Wikipedia, 2023).		Widely used	https://de.wikipedia.org/wiki/Amazon_Luna
7	Arctos Labs Edge Cloud Optimization Solution (ECO)	ECO is a model-driven optimization engine powered by AI. It is intended to add a critical feature to any IT automation system as a plug-in. ECO can understand and make use of the underlying compute grid network metrics as well as application constraints such as latency to match for possible locations, but also impose optimization criteria such as cost for computing and data transport to the final selection. ECO uses an intent-based approach based on expressions of constraints on component interaction, on compute resources, complemented by labels and other properties to determine the best placement. This approach makes ECO capable to consider real-time metrics as well as dynamically adding compute locations to conduct a holistically optimal placement (Arctos labs, 2023).		Used in demos	https://www.arctoslabs.com/solution/

Funded by Horizon Europe Framework Programme of the European Union

8	Avassa Edge Enforcer	The Avassa Edge Enforcer is a container application installed on each host, enabling unique edge site autonomy and application continuity — while having a uniquely small performance footprint. Avassa Edge Enforcer is a key differentiator compared to many other over-the-top or fleet management solutions in that it removes the burden of managing a complex software stack on the edge sites. The Control Tower performs automatic upgrades of the Edge Enforcer. The Edge Enforcer provides cluster-local event streaming and secrets management. Secrets and events are autonomously maintained on the site. The events can be distributed to your central solution. Secrets, in contrast, are configured centrally and securely distributed to the sites ('Avassa Edge Platform', 2022).		Already used and marketed	https://avassa.io/platform/#edge-enforcer
9	BarMan: A run-time management framework in the resource continuum	This work presents and evaluates the BarMan open-source framework by implementing a Fog video surveillance use-case. BarMan leverages a task-based programming model combined with a run-time resource manager and the novel BeeR framework to deploy the application's tasks transparently. This enables the possibility of considering aspects related to the energy and power dissipated by the devices and the single application. Moreover, we developed a task allocation policy to maximize application performance, considering run-time aspects, such as load and connectivity, of the time-varying available devices (Zanella, Sciamanna, & Fornaciari, 2022).	open	Idea/study/ Research project	https://www.sciencedirect.com/science/article/abs/pii/S2210537922000075
10	BOINC	BOINC lets you help cutting-edge science research using your computer. The BOINC app, running on your computer, downloads scientific computing jobs and runs them invisibly in the background. It's easy and safe. About 30 science projects use BOINC. They investigate diseases, study climate change, discover pulsars, and do many other types of scientific research. The BOINC and Science United projects are located at the University of California, Berkeley and are supported by the National Science Foundation (BOINC Berkeley, 2023).	https://github.com/BOINC/boinc	Widely used	https://boinc.berkeley.edu/
11	EdgeX foundry	EdgeX Foundry focus is to exploit the benefits of edge Computing by leveraging cloud-native principles (e.g. loosely-coupled microservices, platform-independence) and by enabling an architecture that meets specific needs of the IoT edge including different connectivity protocols, security and system management for widely distributed compute nodes, and scaling down to highly-constrained devices. The project's sweet spot is in use cases where local decisions are at/or near real-time and when automation and actuation are supported by multiple sources of edge data. Here EdgeX addresses critical interoperability challenges for edge nodes and edge data normalisation where "south meets north, east, and west" in a distributed IoT edge architecture (The Linux Foundation, 2023a).	https://github.com/edgexfoundry	First use cases	https://www.edgexfoundry.org/
12	ExtremeCloud™ Edge	ExtremeCloud Edge gives enterprises more options and greater flexibility to run network applications from any location - including management, analytics, and artificial intelligence (AI). ExtremeCloud Edge is a solution for public, private and edge cloud deployments that reduces the complexity and operational costs of managing cloud applications. Extreme now offers customers the flexibility to create the network edge whenever and wherever they want (Extreme Networks, 2023).		Available for selected partners	https://de.extremenetworks.com/pressemeldungen/press-release/extreme-entwickelt-erstes-networking-cloud-

Funded by Horizon Europe Framework Programme of the European Union

					continuum-fuer-mehr-flexibilitaet-und-cloud-souveraenitaet/
13	FLEDGE	Fledge is an open source framework and community for the industrial edge focused on critical operations, predictive maintenance, situational awareness and safety. Fledge is architected to integrate Industrial Internet of Things (IIoT), sensors and modern machines with the cloud and existing "brown field" systems like historians, DCS (Distributed Control Systems), PLC (Program Logic Controllers) and SCADA (Supervisory Control and Data Acquisition). All sharing a common set of administration and application APIs. Fledge developers and operators no longer face complexity and fragmentation issues when building their IIoT applications as they gather and process more sensor data to automate and transform business. Fledge's modern pluggable architecture eliminates the data silos often found in plants, factories and mines. By using a consistent set of RESTful APIs to develop, manage and secure IIoT applications, Fledge creates a unified solution (The Linux Foundation, 2023b).	https://github.com/fledge-iot	First use cases	https://www.lfedge.org/projects/fledge/
14	Google Cloud Platform	Google Cloud empowers you to quickly build new apps and modernize existing ones to increase your agility and reap the benefits of the multicloud. We offer a consistent platform and data analysis for your deployments no matter where they reside, along with a service-centric view of all your environments (Google Cloud, 2023).		Widely used	https://cloud.google.com/multicloud/
15	Great bear edge by Cisco	Bringing the Cloud-Native Development and Operations Experience to the Edge: We are developing a SaaS-based solution to simplify the deployment and operation of next-generation applications at the edge across large numbers of sites, for developers and IT users, supported with intuitive SDKs and ready-to-use seed applications. This solution optimizes edge applications, data security and real-time experience acceleration. Great Bear is the premier Edge-Native platform that simplifies the development, deployment, and operation of next-generation apps at scale. One platform for all your applications (great bear, 2023).		Already first use cases	https://www.greatbear.app/product
16	ICOS	The unstoppable proliferation of novel computing and sensing device technologies, and the ever-growing demand for data-intensive applications in the edge and cloud, are driving the next wave of transformation in computing systems architecture. The resulting paradigm shift in computing is centred around dynamic, intelligent and yet seamless interconnection of IoT, Edge and Cloud resources in one computing system, to form a continuum (ICOS, 2023).	open	Research phase (parallel research project to FLUIDOS)	icos-project.eu
17	Incode	Innovating and creating a wide-open, secure and trusted IoT-to-edge- to-cloud compute continuum to realize the true potential of edge intelligence (Incode, 2023).	open	Research phase	incode-project.eu

Funded by Horizon Europe Framework Programme of the European Union

18	Integrated Edge Management Model (Dell Technologies)	IEM helps with managing and orchestrating a set of edge capabilities to: - Deploy edge computing that can scale when onboarding new use cases - Bring transparency and trustworthiness to enterprise leadership with a collective understanding of performance metrics across multiple factories - Support the governance and management of multiple applications and compute infrastructure from the near edge or cloud - Streamline technology management (IT and OT) for optimising business and operational activities at each facility - Help reduce costs with multicloud strategies for reliable workload task allocations and, as a result, improve performance of real-time applications - Prevent single point of failures and workload location biases with the adaptation of distributed computing - Contribute to the standardisation of best practices and guidelines across facilities for deploying edge compute infrastructure, thus reducing support costs - Drive the individual facility path to smart manufacturing based on its edge ecosystem (Dell technologies, 2023).		idea	https://infohub.delltechnologies.com/l/integrated-edge-management-in-smart-manufacturing-white-paper-1/summary-910
19	Istio	Istio is an open source service mesh that layers transparently onto existing distributed applications. Istio's powerful features provide a uniform and more efficient way to secure, connect, and monitor services. Istio is the path to load balancing, service-to-service authentication, and monitoring – with few or no service code changes. Its powerful control plane brings vital features, including: · Secure service-to-service communication in a cluster with TLS encryption, strong identity-based authentication and authorization · Automatic load balancing for HTTP, gRPC, WebSocket, and TCP traffic · Fine-grained control of traffic behaviour with rich routing rules, retries, failovers, and fault injection · A pluggable policy layer and configuration API supporting access controls, rate limits and quotas · Automatic metrics, logs, and traces for all traffic within a cluster, including cluster ingress and egress (Istio, 2023).	https://github.com/istio/istio	Already in use	https://istio.io/latest/about/service-mesh/ https://istio.io/latest/news/releases/0.x/announcing-0.1/
20	Kairos	With Kairos you can build immutable, bootable Kubernetes and OS images for your edge devices as easily as writing a Dockerfile. Optional P2P mesh with distributed ledger automates node bootstrapping and coordination. Updating nodes is as easy as CI/CD: push a new image to your container registry and let secure, risk-free A/B atomic upgrades do the rest (The Linux Foundation, 2023c).	https://github.com/kairos-io/kairos	Test phase	https://kairos.io/
21	Karmada	Karmada (Kubernetes Armada) is a Kubernetes management system that enables you to run your cloud-native applications across multiple Kubernetes clusters and clouds, with no changes to your applications. By speaking Kubernetes-native APIs and providing advanced scheduling capabilities, Karmada enables truly open, multi-cloud Kubernetes.	https://github.com/karmada-io/karmada	First use cases	https://karmada.io/

Funded by Horizon Europe Framework Programme of the European Union

		Karmada aims to provide turnkey automation for multi-cluster application management in multi-cloud and hybrid cloud scenarios, with key features such as centralized multi-cloud management, high availability, failure recovery, and traffic scheduling (Karmada, 2023).			
22	KubeEdge	KubeEdge is an open source system for extending native containerized application orchestration capabilities to hosts at Edge. It is built upon kubernetes and provides fundamental infrastructure support for network, application deployment and metadata synchronization between cloud and edge. KubeEdge is licensed under Apache 2.0. and free for personal or commercial use absolutely. We welcome contributors! Our goal is to make an open platform to enable Edge computing, extending native containerized application orchestration capabilities to hosts at Edge, which is built upon kubernetes and provides fundamental infrastructure support for network, app deployment and metadata synchronization between cloud and edge (KubeEdge, 2023).	https://github.com/kubeedge/kubeedge	Test phase	https://kubeedge.io/
23	Limebird – edge it	The decentralized platform enables a self organising and resilient system. An integrated energy management system allows for cross-location optimization (limebird, 2023).		Already in use	https://www.edge-it.io/
24	Liqo	Liqo is an open-source project that enables dynamic and seamless Kubernetes multi-cluster topologies, supporting heterogeneous on-premise, cloud and edge infrastructures (Liqo, 2023).	https://github.com/liqotech/liqo	Test phase	https://liqo.io/
25	Microsoft Azure	Cloud intelligence deployed locally on IoT edge devices Deploy Azure IoT Edge on premises to break up data silos and consolidate operational data at scale in the Azure Cloud. Remotely and securely deploy and manage cloud-native workloads—such as AI, Azure services, or your own business logic—to run directly on your IoT devices. Optimize cloud spend and enable your devices to react faster to local changes and operate reliably even in extended offline periods (Microsoft Azure, 2023).		Widely used	https://azure.microsoft.com/en-us/products/iot-edge/
26	MQTT	MQTT (Message Queuing Telemetry Transport) is a messaging protocol for constrained low-bandwidth networks and extremely high-latency IoT devices. Since Message Queuing Telemetry Transport is specialized for low-bandwidth, high-latency environments, it is an ideal protocol for machine-to-machine (M2M) communication. MQTT works on the publisher / subscriber principle and is operated via a central broker. This means that the sender and receiver have no direct connection. The data sources report their data via a publish and all recipients with interest in certain messages ("identified by the topic") receive the data because they have registered as subscribers (OPC router, 2022).		Already in use	https://www.opc-router.com/what-is-mqtt/

Funded by Horizon Europe Framework Programme of the European Union

27	NebulOus	NebulOuS will develop a novel Meta Operating System and platform for enabling transient fog brokerage ecosystems that seamlessly exploit edge and fog nodes, in conjunction with multi-cloud resources, to cope with the requirements posed by low latency applications. NebulOuS will accomplish substantial research contributions in the realms of cloud and fog computing brokerage by introducing advanced methods and tools for enabling secure and optimal application provisioning and reconfiguration over the cloud computing continuum (Nebulous, 2023).	open	Research phase (parallel research project to FLUIDOS)	nebulouscloud.eu
28	NEMO	Introducing an open source, modular and cybersecure meta operating system, leveraging existing technologies and introducing novel concepts, methods, tools, testing and engagement campaigns (NEMO, 2023).	open	Research phase (parallel research project to FLUIDOS)	meta-os.eu
29	Nephele	Enable efficient, reliable and secure end-to-end orchestration of hyper-distributed applications over programmable infrastructure that spans across the compute Cloud-Edge-IoT continuum, removing existing openness and interoperability barriers in the convergence of IoT technologies against cloud and edge computing orchestration platforms, and introducing automation and decentralized intelligence mechanisms powered by 5G and distributed AI technologies (nephele, 2023).	open	Research phase (parallel research project to FLUIDOS)	nephele-project.eu
30	Oasees	The OASEES project will deliver a European, fully open-source, decentralized, and secure Swarm programmability framework for edge devices and leveraging various AI/ML accelerators (FPGAs, SNNs, Quantum) while supporting a privacy-preserving Object ID federation process (OASEES, 2023).	open	Research project (framework)	oasees-project.eu
31	OpenSwarm	OpenSwarm aims to trigger the next revolution in these data-driven systems by developing true collaborative and distributed smart nodes, through groundbreaking R&I in three technological pillars: efficient networking and management of smart nodes, collaborative energy-aware Artificial Intelligence (AI), and energy-aware swarm programming (OpenSwarm, 2023).	open	Research phase	openswarm.eu
32	Orbital Edge Computing	We propose an Orbital Edge Computing (OEC) architecture to address the limitations of a bent-pipe architecture. OEC supports edge computing at each camera-equipped nanosatellite so that sensed data may be processed locally when downlinking is not possible. In order to address edge processing latencies, OEC systems organize satellite constellations into computational pipelines. These pipelines parallelize both data collection and data processing based on geographic location and without the need for cross-link coordination (Denby & Lucia, 2020).		Idea/study/ research phase	https://dl.acm.org/doi/10.1145/3373376.3378473

Funded by Horizon Europe Framework Programme of the European Union

33	Spectro cloud: CNCF Cluster API	Latest contribution to the open source community allows any team to use Cloud Native Computing Foundation's (CNCF) Cluster API to easily provision and manage Kubernetes clusters on bare metal servers, leveraging Canonical's Metal-as-a-Service (MaaS) interface. This latest open source contribution from Spectro Cloud is the first industry solution to bridge two paradigm shifts for Kubernetes: Cluster API, which enables comprehensive lifecycle management in a declarative manner for Kubernetes environments, and bare metal performance. It allows any team to use CNCF's Cluster API to easily provision Kubernetes clusters on bare metal servers in any data centre and manage them just like any other Kubernetes deployment target, without the extra hassle and risks, or additional technical skill sets. The open source software, under Apache licence and available for anyone to download, can be used by IT and development teams that are interested in managing Kubernetes environments at scale and have or are interested in extending to bare metal servers (spectro cloud, 2023).	https://github.com/spectrocloud	Already in use	https://www.spectrocloud.com/solutions/edge
34	TaRDIS	TaRDIS addresses the ever-increasing complexities of developing correct and performant heterogeneous swarms – swarm systems that are heterogeneous, intelligent, dynamic, and decentralized – by providing a novel programming model, integrated development and analysis environment, and corresponding runtime support (Tardis, 2023).		Research phase	project-tardis.eu

Funded by Horizon Europe Framework Programme of the European Union

4.2 DIFFERENTIATION OF FLUIDOS COMPARED TO EXISTING TECHNOLOGIES WITH TECHNOLOGY RADARS

The following comparison juxtaposes FLUIDOS against alternative technical approaches and market offerings. For this purpose, technology radars serve as graphical tools employing two axes: one axis classifies the existing level of technical maturity or market adoption (vertical), while the other axis assesses a technical attribute or feature (horizontal). Two characteristics were chosen in advance to differentiate the distributed computing approaches found in terms of their character:

- First, the extent of how distributed the compute capacity will be provided (extreme cases: exclusively in the cloud vs. exclusively on end devices).

- Second, whether they are aimed at specific applications or at a (possibly generic) multipurpose computing environment.

Generally, only few of the solutions are already represented on a larger market. The latter tend to be from larger players in the cloud computing market (see AWS, Microsoft, Google) and therefore generate a lot of attention, but in most cases only offer proprietary approaches.

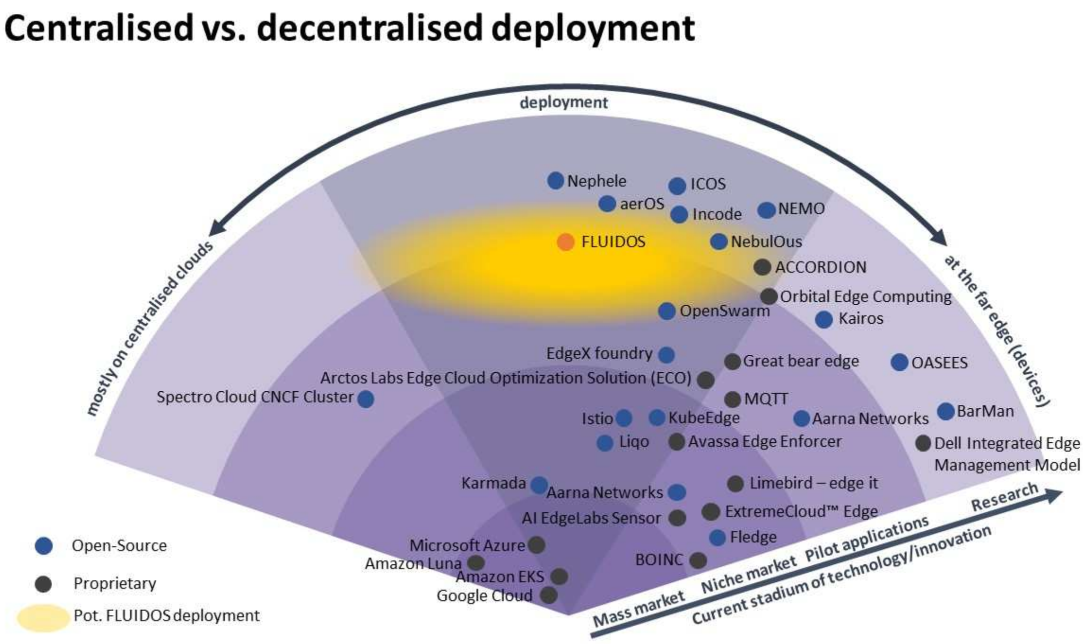

Figure 4: Technology radar – centralized vs. decentralized deployment of approaches

The second distinction, on the other hand, focuses on whether the (possibly edge) computing environments are being developed for specific applications or application domains, or whether a very general, multipurpose edge cloud environment should be enabled. Many of the existing (cloud) computing environments are naturally designed to fit for various applications (see Google Cloud, Amazon EC2 etc.).

Funded by Horizon Europe Framework Programme of the European Union

All the applications that are classified in a similar realm between cloud and edge (cloud edge continuum) are often multipurpose focused (see EdgeX foundry, ICOS). Many of the latest edge approaches (decentralized) that are being developed in the last few months or years are like FLUIDOS available open source. Some of them are focussed on an application-specific approach. Fledge, for example, provides a platform for industrial computing environments at the edge and Orbital Edge Computing focusses on satellite image processing. Approaches already on the market typically focus on a specific purpose and/or are not open source.

An intent-driven technology similar to FLUIDOS, which offers a dynamic allocation between different deployment models depending on the user's need/use case, that is already widely in use, has not been identified among the currently available offerings. One of the most relevant approaches is the “Arctos Labs Edge Cloud Optimization Solution”. Multiple other research projects on EU level on MetaOS (aerOS, iCOS, NebulOuS, NEMO, Nephele) which receive funding from the same call follow similar approaches as FLUIDOS.

Particularly for this rationale, and also due to the heightened transparency regarding the capabilities of various clusters or nodes, FLUIDOS possesses the ability to amalgamate the benefits inherent in both realms in this context. While we would classify it more on the multipurpose side, it also offers the possibility to select a specific device when required for a specific purpose of computing task. This flexibility, lacking in other approaches to date, also makes the use of resources/energy much more efficient. Meanwhile, the privacy of both the end user and the host of the services is maintained.

More established projects like KubeEdge also aimed at connecting Cloud and Edge could provide a good lead/model for FLUIDOS in some areas. KubeEdge lacks the built-in communication and network, which would be an aspect that FLUIDOS is striving for.

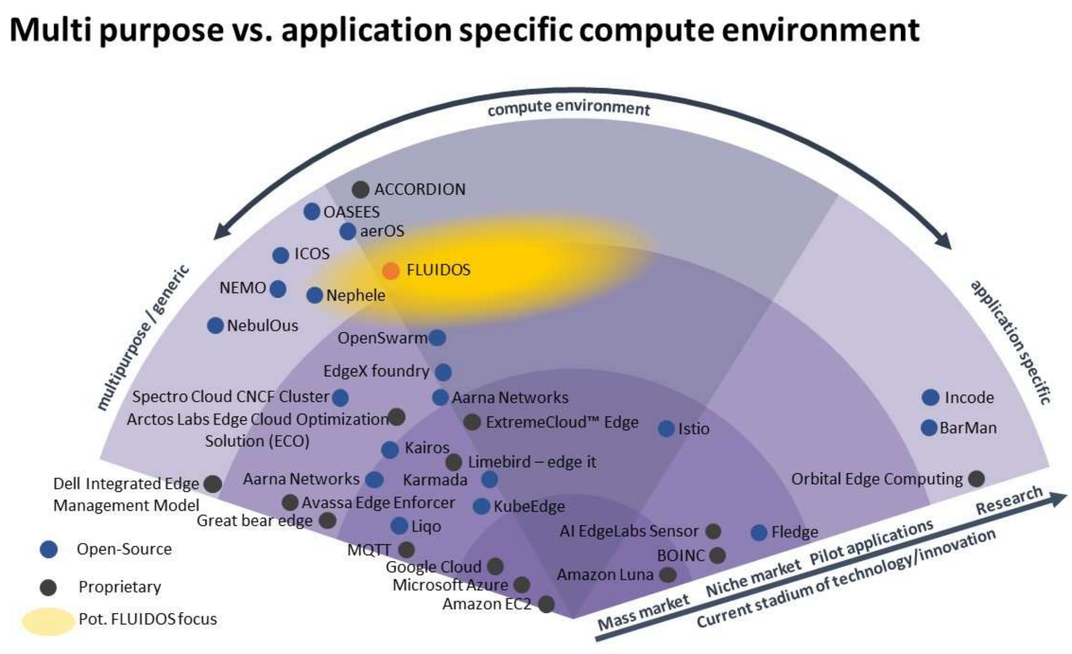

Figure 5: Technology radar - edge(-cloud) approaches regarding multipurpose vs. application specific compute environments

Funded by Horizon Europe Framework Programme of the European Union

4.3 IMPORTANT CHARACTERISTICS AND FEATURES OF FLUIDOS

Some conclusions can be drawn from the interviews as to which features would be important or even essential for different potential user groups of FLUIDOS:

FLUIDOS should be a compact and easy to use software, which provides all important basic functions. If the software is overloaded with features and settings and is difficult to understand or use, it will have little chance of success. Additional features should be able to be added incrementally, because depending on the use case and requirements, completely different functions and software components may be needed, which should not automatically be integrated for all users to reduce complexity.

Security and the zero-trust paradigm are essential components of FLUIDOS, indispensable for many users. At the same time, FLUIDOS is intended to run software on various devices with potentially self-protecting data, which in turn processes sensitive data - both the device operator and the edge cloud application must have absolute trust in the paradigm, as mistrust would destroy any benefits. This requires not only an absolutely reliable implementation of the Zero Trust paradigm, but also a comprehensible and convincing description - both for the user/owner of the devices and for the operator of the edge cloud application.

It shall also be considered whether and to what extent the power requirements of the host devices increase, as this factor can have a significant impact on the usability of battery powered devices in particular. A software-integrated policy that automatically adjusts the release of computing resources in the network according to the energy status (depending on plug/battery/battery level) would probably be an essential component.

Reliability was another very significant issue raised in the interviews. In sectors as diverse as manufacturing, power grids and hospitals, it is vital that IT systems are reliable. FLUIDOS must at least match the standards of established systems in order to have a chance of playing a role there. If FLUIDOS also has the ability to automatically and dynamically offload IT workloads to other devices or to a cloud in the event of a failure or power outage, this could even become an important motivation to replace older software with FLUIDOS or to run on a FLUIDOS based platform.

Funded by Horizon Europe
Framework Programme of the European Union

5 Requirements for a wide adoption and exploitation of the project results

FLUIDOS is still at a relatively early technological stage in the research project; as already described, the aim is to reach TRL 5 by the end of the project. This means that the technology should already have been tested in a relevant environment. On the one hand, this should be done comprehensively via the use cases planned in the project in Work Package 7 as well as the open calls via Work Package 8 by the end of the project. This should mean that, at the end of the project, a critical number of potential users will already have been exposed to FLUIDOS, providing a very favourable starting point for further dissemination and application.

With its unique characteristics of a zero trust, fully automated, intent driven and ecological cloud edge continuum, FLUIDOS fills a niche with enormous potential for the future cloud computing market. Further development to higher TRLs on the one hand and market penetration and high awareness in the open source community on the other hand are interdependent. In order to make use of the results and foundations created in the FLUIDOS research project, some early recommendations can already be given today. In other work packages of FLUIDOS, much more detailed work plan and exploitation strategy will be worked out. The early recommendations from this document are:

1. A critical number of software developers and software companies need to continue working/building on the idea of FLUIDOS after the project ends. This is essential so that the technology does not end before it is deployed, as is often the case in an early stage of development. On the one hand, it is of great importance that the level of awareness of FLUIDOS is increased. In addition to potential users, it is important to reach a broader developer community. To achieve this, it is essential to have good documentation of the software and its operation as well as to publish the code of FLUIDOS, which is currently already in progress on Github (GITHUB, 2023). The source code is published under the open source licence Apache 2.0. This licence allows for both non-commercial and commercial use of the software. This flexibility can encourage wider adoption of the software in various contexts, including business and commercial projects. Unlike some other open-source licences (e.g., GNU GPL), the Apache Licence 2.0 does not have a "viral" or "copyleft" clause that requires derivative works to also be open source. This means that developers can combine Apache-licensed code with code under other licences, including proprietary licences, without requiring the entire project to be open source. The Apache Licence 2.0 includes a patent grant clause, which provides some assurance to users and contributors that they won't be subject to patent-related legal action from the project's contributors. This can help create a more predictable and stable environment for collaboration. The Apache Software Foundation (ASF) often uses the Apache Licence 2.0 for its projects, fostering a sense of community and collaboration around software development. Many developers are familiar with this licence and may be more inclined to participate in projects that deploy it. All these features of documentation, publication and open-source licensing thus already provide a good basis for further dissemination.
2. Raise awareness of FLUIDOS within possible user communities: Already during the project and afterwards, extensive public relations work is necessary to reach the expert community as well as potential users or other interested parties and software developers. On the one hand, this requires a solid and up-to-date website, which already exists to a certain extent (https://www.fluidos.eu), but which needs to be further expanded with current and up-to-date publications and news about FLUIDOS. In addition, a continuous presence in professional

Funded by Horizon Europe Framework Programme of the European Union

journals and scientific conferences as well as in associations and communities such as the Cloud Native Computing Foundation is necessary. An impactful presence in social media would also be desirable, although experience shows that FLUIDOS can be much more successful if the project partners themselves spread the idea of FLUIDOS, rather than exclusively via new accounts with limited reach. By project partners, we mean on the one hand the institutions involved in FLUIDOS, but even more the real people who have their own personal networks. A fundamental difficulty will be to continue this external communication after the end of the project. To this end, a strategy should already be developed during the funded project period as to who in general will continue to work on FLUIDOS afterwards (possibly without funding) and how, for example, access to social media accounts etc. will be handled.

3. <u>Key Partners for further dissemination:</u> During the interviews, it was often mentioned that a project as young as FLUIDOS is necessarily dependent on further partnerships. The already existing participation in the EUCloudEdgeIoT.eu initiative offers many networking opportunities with similar, partly competing approaches, but also with partners that are upstream or downstream in the digital value chain. Here you can find several possible partners who are on the same scale/level; possibly also for future research projects. In addition, the existing cloud providers can of course become a very important key partner for FLUIDOS in the future. As described above, most cloud providers already have edge cloud projects underway, but FLUIDOS could provide an independent solution that allows even more dynamic (fluid) movement of workloads, even across providers. FLUIDOS could also support European and possibly smaller cloud providers, making it easier for them to play a role in a market that is still very much driven by economies of scale. In order to get existing top dogs interested in FLUIDOS at all, it is significant to promote FLUIDOS to their potential European customers, e.g. SMEs, public administrations in Europe and innovative start-ups. The open calls in Work Package 8 could also provide an essential basis for this.
4. <u>Find attractive niches to launch FLUIDOS and then scale up:</u> There is no doubt that there is already immense potential for unused computing and storage capacity in edge devices. This capacity is underutilized, mainly because it is small and distributed. The development of always-on, always-available networks and the ability of virtualization/cloud to run applications in a highly distributed manner are both factors that strongly suggest that this capacity will be used sooner or later. In addition, the increasing resource and energy consumption of the cloud is being discussed more and more critically. FLUIDOS is an enabling technology that aims to unlock this huge potential with a promising open source approach. In order to prevail against other competitors who have recognized the same potential, it is significant to have a foot in the door of key users (early adopters) at an early stage. These could be the sectors of the use cases represented in the project, but possibly also others, such as the medical sector. Choosing the right niche in the market is essential to implementing the benefits of FLUIDOS in the software landscape and to communicate them to a wider audience. At the same time, it is also necessary to prepare for strong scaling, as the history of digitization shows the significance of scale and network effects for the market penetration of applications. One enticing possibility would be to integrate FLUIDOS into a Linux distribution as a scheduler, where the release of unused computing capacity for network users or similar could be activated with a few mouse clicks. Within organizations or for non-profit purposes, this can be organized free of charge; beyond that, remuneration through pay-per-use business models is also conceivable.

Funded by Horizon Europe Framework Programme of the European Union

6. LITERATURE

Abbas, N., Zhang, Y., Taherkordi, A., & Skeie, T. (2017). Mobile edge computing: A survey. IEEE Internet of Things Journal, 5(1), 450–465.

Aeros. (2023). Aeros-project.eu – aerOS project. Retrieved 6 July 2023, from https://aeros-project.eu/

Arctos labs. (2023). ECO Solution. Retrieved 6 July 2023, from Arctos Labs—Optimization for edge computing website: https://www.arctoslabs.com/solution/

Ars technica. (2023, July 11). Is distributed computing dying, or just fading into the background? Retrieved 23 August 2023, from Ars Technica website: https://arstechnica.com/science/2023/07/is-distributed-computing-dying-or-just-fading-into-the-backdrop/

Avassa Edge Platform. (2022). Retrieved 6 July 2023, from https://avassa.io/platform/

AWS. (2023). Secure and resizable cloud compute – Amazon EC2 – Amazon Web Services. Retrieved 27 July 2023, from Amazon Web Services, Inc. website: https://aws.amazon.com/ec2/

Bittman, T. (2017, March 6). The Edge Will Eat The Cloud. Retrieved 25 October 2017, from Gartner Blog Network website: https://blogs.gartner.com/thomas_bittman/2017/03/06/the-edge-will-eat-the-cloud/

BOINC Berkeley. (2023). BOINC. Retrieved 14 July 2023, from https://boinc.berkeley.edu/

Bonomi, F., Poledna, S., & Steiner, W. (2017). The role of fog computing in the future of the automobile. *Fog for 5G and IoT*, *191*.

Cao, K., Liu, Y., Meng, G., & Sun, Q. (2020). An overview on edge computing research. *IEEE Access*, *8*, 85714–85728.

Cutress, D. I. (2020). Insights into DDR5 Sub-timings and Latencies. Retrieved 27 September 2023, from https://www.anandtech.com/show/16143/insights-into-ddr5-subtimings-and-latencies

DataCenter Insider. (2021). Embedded-Servermodule für Edge-Rechenzentren. Retrieved 23 August

Funded by Horizon Europe Framework Programme of the European Union

2023, from https://www.datacenter-insider.de/embedded-servermodule-fuer-edge-rechenzentren-a-85af3b1d9cd3e1b94965383d56d6e96d/

Dell technologies. (2023). Summary | Integrated Edge Management in Smart Manufacturing—White Paper | Dell Technologies Info Hub. Retrieved 6 July 2023, from https://infohub.delltechnologies.com/l/integrated-edge-management-in-smart-manufacturing-white-paper-1/summary-910

Denby, B., & Lucia, B. M. (2020). Orbital Edge Computing | Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems. Retrieved 6 July 2023, from https://dl.acm.org/doi/10.1145/3373376.3378473

Edge Industry Review. (2023). Aarna Networks launches AES, a zero-touch orchestration as a service for edge computing | Edge Industry Review. Retrieved 6 July 2023, from https://www.edgeir.com/aarna-networks-launches-aes-a-zero-touch-orchestration-as-a-service-for-edge-computing-20230226

Edge labs. (2023). AI Edge labs. Retrieved 6 July 2023, from AI EdgeLabs website: http://edgelabs.ai/platform/

Extreme Networks. (2023). Extreme entwickelt erstes Networking Cloud Continuum für mehr Flexibilität und Cloud-Souveränität. Retrieved 6 July 2023, from Extreme Networks Germany website: https://de.extremenetworks.com/?post_type=news_item&p=4868

Gartner. (2018). What Edge Computing Means For Infrastructure And Operations Leaders. Retrieved 23 August 2023, from https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders

GITHUB. (2023). fluidos-project/Docs: This repository aims to be the single source of truth about the architecture and the behavior of FLUIDOS. It will also be an indexing point for all the artifacts and repositories involved. Retrieved 27 September 2023, from https://github.com/fluidos-

Funded by Horizon Europe Framework Programme of the European Union

project/Docs

Global Market Insights. (2020). Edge Data Center Market Share 2020-2026 | Global Report. Retrieved 21 October 2020, from Global Market Insights, Inc. website: https://www.gminsights.com/industry-analysis/edge-data-center-market

Global Market Insights. (2022). Edge Data Center Market Size & Share | Global Report 2023-2032. Retrieved 23 August 2023, from Global Market Insights Inc. website: https://www.gminsights.com/industry-analysis/edge-data-center-market

Google Cloud. (2023). Multi-Cloud-Lösungen | Google Cloud. Retrieved 14 July 2023, from https://cloud.google.com/multicloud?hl=de

Grand View Research. (2022). Edge Computing Market Size, Share & Growth Report, 2030. Retrieved 23 August 2023, from https://www.grandviewresearch.com/industry-analysis/edge-computing-market

great bear. (2023). Great bear | Product. Retrieved 6 July 2023, from https://www.greatbear.app/product

Hintemann, R., Hinterholzer, S., & Grothey, T. (2022). Nachhaltiges Edge Computing im Kontext der Machine Economy. In D. Wurm (Ed.), *Wege in eine ökologische Machine Economy Ein systemischer Blick auf die Umweltwirkungen digitaler Technologien.* (pp. 170–193). Berlin, Wuppertal: Institut für Zukunftsstudien und Technologiebewertung (IZT) & Wuppertal Institut für Klima, Umwelt, Energie.

Hintemann, R., Hinterholzer, S., Montevecchi, F., & Stickler, T. (2020). *Energy-efficient Cloud Computing Technologies and Policies for an Eco-friendly Cloud Market*. Berlin, Vienna: Borderstep Institute & Environment Agency Austria. Retrieved from Borderstep Institute & Environment Agency Austria website: https://op.europa.eu/en/publication-detail/-/publication/bf276684-32bd-11eb-b27b-01aa75ed71a1/language-en/format-PDF/source-183168542

Funded by Horizon Europe Framework Programme of the European Union

ICOS. (2023). IoT to Cloud Operating System | ICOS. Retrieved 6 July 2023, from https://www.icos-project.eu/

IDC Corporate. (2023). New IDC Spending Guide Forecasts Edge Computing Investments. Retrieved 23 August 2023, from IDC: The premier global market intelligence company website: https://www.idc.com/getdoc.jsp?containerId=prUS50386323

Incode. (2023). Homepage. Retrieved 6 July 2023, from INCODE project website: https://www.incode-project.eu/

Internet Archive. (2023). Retrieved 21 September 2023, from https://archive.org/web/

Istio. (2023). Introducing Istio. Retrieved 6 July 2023, from Istio website: https://istio.io/latest/news/releases/0.x/announcing-0.1/

JM Chabas, Chandra Gnanasambandam, Sanchi Gupte, & Mitra Mahdavian. (2018). What edge computing means for hardware companies | McKinsey. Retrieved 17 January 2022, from https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/new-demand-new-markets-what-edge-computing-means-for-hardware-companies

Karmada. (2023). Open, Multi-Cloud, Multi-Cluster Kubernetes Orchestration | karmada. Retrieved 24 August 2023, from https://karmada.io/

Khan, L. U., Yaqoob, I., Tran, N. H., Kazmi, S. A., Dang, T. N., & Hong, C. S. (2020). Edge-computing-enabled smart cities: A comprehensive survey. *IEEE Internet of Things Journal*, *7*(10), 10200–10232.

KubeEdge. (2023). KubeEdge. Retrieved 24 August 2023, from Welcome to KubeEdge website: https://kubeedge.io/

limebird. (2023). LIMEBIRD – Green and distributed cloud computing. Retrieved 6 July 2023, from https://www.edge-it.io/

Liqo. (2023). Liqo.io. Retrieved 24 August 2023, from Enable dynamic and seamless Kubernetes multi-cluster topologies website: https://liqo.io/

Funded by Horizon Europe Framework Programme of the European Union

Marshall, P. (2021). *STATE OF THE EDGE 2021—A Market and Ecosystem Report for Edge Computing*. online: State of the edge. Retrieved from State of the edge website: https://stateoftheedge.com/reports/state-of-the-edge-report-2021/

Marshall, P. (2022). *STATE OF THE EDGE 2022—The New Open "Edge"*. online: State of the edge. Retrieved from State of the edge website: https://stateoftheedge.com/reports/state-of-the-edge-report-2022/

Microsoft Azure. (2023). IoT Edge | Cloud Intelligence | Microsoft Azure. Retrieved 14 July 2023, from https://azure.microsoft.com/de-de/products/iot-edge

Nebulous. (2023). NebulOuS Meta Operating System and platform. Retrieved 6 July 2023, from https://www.nebulouscloud.eu/

NEMO. (2023). HOME - META-OS. Retrieved 6 July 2023, from https://meta-os.eu/

nephele. (2023). A lightweight software stack and synergetic meta-orchestration framework for the next generation compute continuum | Nephele. Retrieved 6 July 2023, from https://nephele-project.eu/

OASEES. (2023). OASEES – Open autonomous programmable cloud apps & smart sensors. Retrieved 6 July 2023, from https://oasees-project.eu/

Omdia. (2021). *Industrial Edge Networking Components – 2020*.

OPC router. (2022). Was ist MQTT? Erklärung mit industriellem Fokus. Retrieved 6 July 2023, from Systemintegration mit dem OPC Router website: https://www.opc-router.de/was-ist-mqtt/

OpenSwarm. (2023). OpenSwarm. Retrieved 6 July 2023, from https://openswarm.eu/

Robert Kahn. (1972). *DEMONSTRATION AT INTERNATIONAL COMPUTER COMMUNICATIONS CONFERENCE (RFC)*. Retrieved from https://www.rfc-editor.org/rfc/rfc371.html

Shi, W., & Dustdar, S. (2016). The promise of edge computing. *Computer*, *49*(5), 78–81.

Shirer, M. (2021). IDC Forecasts Strong Growth for the Managed Edge Services Market with Worldwide Revenues Approaching $2.8 Billion in 2025. Retrieved 4 March 2022, from IDC:

Funded by Horizon Europe Framework Programme of the European Union

The premier global market intelligence company website: https://www.idc.com/getdoc.jsp?containerId=prUS48179321

spectro cloud. (2023). Award-Winning Edge Kubernetes—Spectro Cloud. Retrieved 6 July 2023, from https://www.spectrocloud.com/solutions/edge

Tardis. (2023). Home—TaRDIS project. Retrieved 6 July 2023, from https://www.project-tardis.eu/

The Linux Foundation. (2021). LF Edge's State of the Edge 2021 Report Predicts Global Edge Computing Infrastructure Market to be Worth Up to $800 Billion by 2028—Linux Foundation. Retrieved 23 August 2023, from https://www.linuxfoundation.org/press/press-release/lf-edges-state-of-the-edge-2021-report-predicts-global-edge-computing-infrastructure-market-to-be-worth-up-to-800-billion-by-2028

The Linux Foundation. (2023a). EdgeX Foundry | The Open Source Edge Platform. Retrieved 6 July 2023, from Https://www.edgexfoundry.org website: https://www.edgexfoundry.org

The Linux Foundation. (2023b). Fledge. Retrieved 6 July 2023, from LF Edge website: https://www.lfedge.org/projects/fledge/

The Linux Foundation. (2023c). Kairos—The immutable Linux meta-distribution for edge Kubernetes. Retrieved 6 July 2023, from Kairos—The immutable Linux meta-distribution for edge Kubernetes website: https://kairos.io/

van der Meulen, R. (2019). What Edge Computing Means for Infrastructure and Operations Leaders. Retrieved 12 April 2018, from Smarter With Gartner website: https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders/

Vida Rozite, Emi Bertoli, & Brendan Reidenbach. (2023). Data centres & networks. Retrieved 16 August 2023, from IEA website: https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks

Wikipedia. (2023). Amazon Luna. In *Wikipedia*. Retrieved from

Funded by Horizon Europe Framework Programme of the European Union

https://de.wikipedia.org/w/index.php?title=Amazon_Luna&oldid=232306987

Sustainable Computing: Informatics and Systems, *35*, 100663.

https://doi.org/10.1016/j.suscom.2022.100663

Funded by Horizon Europe Framework Programme of the European Union