

Extending Federated Data Platforms in the Cloud Continuum for Manufacturing and Intralogistics

Krista Mätäsniemi¹^a, Teemu Toroi²^b and David Hästbacka¹^c

¹*Computing Sciences, Faculty of Information Technology and Communication Sciences, Tampere University, 33100 Tampere, Finland*

²*Department of Computer Science, School of Science, Aalto University, 02150 Espoo, Finland*
fi fi

Keywords: Federated Data Platforms, Data Spaces, Cloud Continuum, Service Placement, Edge Deployment, On-premise.

Abstract: Federated data platforms (FDPs), or data spaces, are frameworks that facilitate data sharing between diverse data sources and stakeholders. FDPs are critical enablers in the 2020 European Data Strategy where industrial data is a strategic investment area. Industrial use cases can involve high data volumes and velocity despite operating in resource-constrained environments. FDPs have been proven in inter-company data exchanges and are typically implemented on scalable cloud platforms. This study explores the possibility of extending FDPs in the cloud continuum (CC) by examining how the technical requirements of FDP building blocks align with the drivers, characteristics, and enablers of CC layers. The study introduces an analysis framework for finding the optimal placement of building blocks along CC. It also shows that transferring FDP components connected to high data volumes closer to the edge can make FDPs more effective and better suited for manufacturing and intralogistics use cases. The study concludes that there is an architectural fit between FDP functionalities and the characteristics of the CC, and suggests that FDP usability and performance should be studied empirically.

1 INTRODUCTION


Data is a strategic resource for the optimization of manufacturing operations and the creation of new products and services (Gelhaar et al., 2021) (Richter and Slowinski, 2019). Data must flow between different systems and stakeholders to enable new data-driven innovations and operative efficiency (Rajput and Singh, 2019) (Suleiman et al., 2022).


Transferring data between manufacturing systems is not feasible in several use cases due to large data volumes, latency requirements, transfer costs, communication bandwidth limitations, or security considerations (Alkhabbas et al., 2020) (Sousa et al., 2022). Cloud continuum (CC) can be used alleviate data transfer challenges via deployment of applications and other resources close to data sources and users (Al-Dulaimy et al., 2024) (Cardellini et al., 2025). Placement and scheduling of computing along the CC remain focal architectural and practitioner questions (Jansen et al., 2023) (Salaht et al., 2020).


Federated data platforms (FDPs) are frameworks that facilitate the sharing, integration, and management of data between a diverse network of data sources and stakeholders (Curry et al., 2022) (Otto et al., 2022). FDPs are at the core of the European data strategy, where industrial data sharing is a strategic investment area (European Commission, 2020) (European Commission, 2024). Early FDP research focuses on technical, legal and organizational characteristics (Otto et al., 2016) (Otto et al., 2022), whereas business models and value creation are emerging FDP research themes (Ammann and Hess, 2025) (Jussen et al., 2024). To the best of our knowledge, there is no research addressing how FDPs can be extended in CC and how their components should be placed, i.e. on-premise or as edge deployments.

To address this research gap, our research questions are:

1. How can federated data platforms extend along the cloud continuum?
2. What specific requirements and implications do manufacturing and intralogistics have for the deployment of FDPs along cloud continuum?

^a <https://orcid.org/0009-0008-6627-7156>

^b <https://orcid.org/0009-0005-1617-9943>

^c <https://orcid.org/0000-0001-8442-1248>

We answer the research questions by designing a framework for evaluating architectural fit between FDP building blocks and CC layers, and analyzing the specific requirements of manufacturing and intralogistics for FDP deployment. Both the framework and its application are novel contributions with practitioner impacts.

This paper is structured as follows. Section 2 presents the background and existing literature on FDPs, the cloud continuum and the service placement problem. Section 3 introduces the formulated evaluation framework. Section 4 maps FDP along the CC. Section 5 presents application considerations for manufacturing and intralogistics. Section 6 provides a discussion before conclusions in section 7.

2 THEORETICAL BACKGROUND

2.1 Federated Data Platforms

Federated data platforms are technical, legal, and organizational frameworks that facilitate the sharing, integration, and management of data between a diverse network of data sources and stakeholders (Otto et al., 2016). They provide conceptual, organizational, and technological building blocks for data ecosystems, where stakeholders collaborate via federated data storages (Otto et al., 2022).

As discussed in (Otto, 2022) and (Otto et al., 2022), federation means that FDP architecture and ownership are not centralized but shared by participating actors. Actual data transfers are bilateral and do not require data duplication; instead, data is accessed from its original location. All data in an FDP is cataloged with the FDP operator, and access to it is provided in a decentralized way. The data remain at the source, where each data provider decides to host it. Data is accessed and provided to other participating actors through data connectors. Figure 1 provides a schematic overview of FDP actors and data flows.

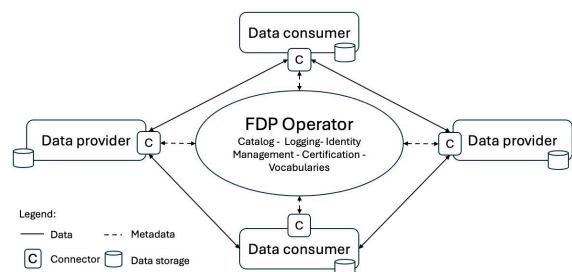


Figure 1: Overview of FDP participant roles participant roles and dataflows in a FDP. Adapted and simplified from (Otto, 2022).

Other key characteristics of FDPs are data sovereignty, interoperability, and scalability (Jarke et al., 2019) (Otto et al., 2022). As discussed by (Jarke et al., 2019), data sovereignty means that data remains under the control of the data provider, who decides who has access to it, for what purpose, and under what conditions their data can be utilized. Sovereignty is enforceable through data contracts attached to the data offering.

FDP frameworks are developed, standardized, and promoted in industry organizations such as the International Data Spaces Association (IDSA), Gaia-X, and the Data Spaces Support Centre (DSSC). DSSC is the most relevant organization for this study, with their Data Spaces Blueprint that defines the business, organizational, and technical building blocks for FDPs (Data Spaces Support Center, 2024).

2.2 Cloud Continuum

CC is a systems-level paradigm related to extending scalable computation, data storage, and other capabilities towards the edge and endpoints of the network (Moreschini et al., 2022) (Bonomi et al., 2014). Key capabilities extended from the cloud towards the edge are computation, storage, application control and communication between network nodes - both horizontally within a network layer and vertically between network layers (Byers and Swanson, 2017).

Industry 4.0, cyber-physical systems and the Internet of Things (IoT) have been major drivers for CC with heterogeneous connected actuators, sensors, and devices producing massive amounts of data with high velocity in industrial contexts (Suleiman et al., 2022) (Lasi et al., 2014). Transferring all that data to the cloud is impractical due to incurred costs, limitations of the communication bandwidth, low latency requirements, and security considerations (Alkhabbas et al., 2020) (Sousa et al., 2022). CC solves these challenges by enabling the deployment of applications and other resources close to data sources and users (Cardellini et al., 2025) (Jansen et al., 2023). There are multiple drivers for extending capabilities closer to edge and endpoints of computing networks. This study categorizes these drivers into business, functional, and security and privacy domains based on earlier research as depicted in Table 1.

Research in CC has been active, and multiple complementary concepts have emerged, such as Edge Computing, Fog Computing, Fog-to-Cloud Computing, Mobile Cloud Computing, Multi-Access Edge Computing, Mist Computing, and the Cloud-to-Thing Continuum (Al-Dulaimy et al., 2024) (Antonini et al., 2019) (Čilić et al., 2021) (Masip-Bruin et al., 2016)

Table 1: Drivers for cloud continuum utilization.

Category	Drivers	References
Business	Cost efficiency Operational reliability and resilience Energy efficiency	(Alkhabbas et al., 2020) (Bellavista et al., 2019) (Cardellini et al., 2025) (Di Martino et al., 2023) (Masip-Bruin et al., 2016) (Moreschini et al., 2022) (Salaht et al., 2020) (Santos et al., 2021) (Sousa et al., 2022) (Trakadas et al., 2022)
Functional	End-to-end performance scalability Low latency requirements High data volumes and velocities Communications (bandwidth efficiency) Interoperability with endpoints and data formats Mobility	(Ahearne et al., 2023) (Alkhabbas et al., 2020) (Barbosa et al., 2025) (Bellavista et al., 2019) (Cardellini et al., 2025) (Di Martino et al., 2023) (Jansen et al., 2023) (Kimovski et al., 2021) (Masip-Bruin et al., 2016) (Moreschini et al., 2022) (Salaht et al., 2020) (Santos et al., 2021) (Sousa et al., 2022) (Trakadas et al., 2022)
Security	Guarantee privacy Avoid or limit transfer of sensitive data Reduce risk surface	(Ahearne et al., 2023) (Alkhabbas et al., 2020) (Cardellini et al., 2025) (Di Martino et al., 2023) (Jansen et al., 2023) (Moreschini et al., 2022) (Santos et al., 2021) (Sousa et al., 2022) (Trakadas et al., 2022)

(Salaht et al., 2020) (Yousefpour et al., 2019). Correspondingly, several related architecture proposals have been published where with 3-8 logical continuum layers. The proposals describe logic, architecture, and procedures how computing, storage, control, and networking capabilities can be organized and distributed along the CC layers; and how applications can be dynamically redistributed. The logical separation into different layers enables architects and developers to consider the continuum as a unified computing model. It helps overcome the limitations of tiered architectures and supports moving applications closer to data sources and end users (Cardellini et al., 2025) (Jansen et al., 2023).

2.3 Service Placement Problem

The Service Placement Problem (SPP), also referred to as the Application Placement Problem (Xiao et al., 2012) or the Function Placement Problem (Jonas et al., 2019), defines a mapping pattern for how application components are assigned to an infrastructure, along with a methodology for how tasks are divided and scheduled across those components (Salaht et al., 2020). Application workloads can be distributed along the infrastructure and leverage the capabilities of each infrastructure layer (Jansen et al., 2023). When the data storage, compute, and communication capabilities of each network element are taken into account, applications can reside close to the data, and significant latency and network congestion issues can be avoided—as large amounts of data are not transferred between network layers but are instead processed in nodes close to data sources (Farhadi et al., 2021) (Poularakis et al., 2020).

The placement pattern can be static or dynamic,

with continuous reallocation and scheduling of resources to achieve a single objective or multiple interdependent objectives simultaneously (Brogi et al., 2020). The optimization objectives discussed in recent SPP research include both functional and non-functional requirements, such as total consumption of network resources (CPU cycles, memory space, data storage, network bandwidth), energy consumption, quality of service (maximum delay, task completion rate and time, service disruption rate), end-user mobility, total costs (application deployment and operations), minimizing data transfer to the cloud, and revenues enabled by the network (Farhadi et al., 2021) (Kasi et al., 2020) (Natesha and Guddeti, 2021) (Ramírez et al., 2017). It is worth noting that many of these requirements align with the business and functional drivers of cloud continuum utilization summarized earlier in Table 1. However, security is not emphasized in most SPP studies, despite being a critical non-functional requirement—particularly regarding the handling and transfer of sensitive and confidential data (Salaht et al., 2020).

For the purposes of this study, it is important to acknowledge that dynamic SPP is a non-trivial problem, even though our focus is limited to defining a static mapping pattern for FDP components along the main CC layers. We do not take a position on data placement options (e.g., (Du et al., 2020) (Shen et al., 2024)) but assume that data is expected to reside close to its source unless there are explicit reasons to transfer it to fog or cloud repositories. This assumption aligns with the design principles of FDPs. By keeping data close to its sources, we avoid problems related to the synchronization of distributed data (Gilbert and Lynch, 2002) (Pandey and Pandey, 2020).

3 ANALYSIS FRAMEWORK

3.1 Characteristics of Cloud Continuum Levels

For the purposes of this study, the CC is divided into three layers. This allows the evaluation of the architectural fit while maintaining a manageable level of complexity. Adopting and modifying the models from (Jansen et al., 2023) and (Salaht et al., 2020), we select the layers cloud, fog, and edge as our analysis baseline. Each layer has its own set of characteristics that present both benefits and challenges for application deployments. These benefits and challenges are summarized in Table 2.

The authors of (Chiang and Zhang, 2016) describe the cloud as a small number of very large data centers that offer a centralized environment for computing, control, and data storage. Data centers are operated by large companies and maintained by technical expert teams ensuring high operational reliability and the security of their services. When the majority of security aspects are the responsibility of the cloud service provider, deploying a system in the cloud enhances the overall security of the system and reduces the security-related risk borne by the system provider or owner (Alkhabbas et al., 2020). However, a cloud-centric approach presents two major limitations regarding system privacy and reliability (Antonini et al., 2019). Firstly, sensitive data should not be transmitted or processed in remote systems due to increased cybersecurity risks. Secondly, securing the reliability of the system requires an always-on Internet connection, which can be difficult on mobile devices.

In addition to security and reliability, cloud infrastructure provides scalable and extensive computational resources and storage capabilities to enable high performance processing and resource-intensive applications, such as deep learning (Jansen et al., 2023). IoT applications create major challenges for cloud-centric systems due to the high volumes of data and the increased number of connected devices. Transferring large amounts of data to the cloud for processing has a high bandwidth demand and consumes energy (Alkhabbas et al., 2020). Moreover, the billing models of cloud services are typically based on the number of transactions with the cloud, making continuous streaming of data to the cloud costly (Antonini et al., 2019). Lastly, the geographical distance between devices and the cloud introduces delays in communication which makes cloud-centric systems unsuitable to latency-sensitive applications (Barbosa et al., 2025).

Fog acts as an intermediate layer between the cloud and end devices, bringing computing, control, storage, and networking functions closer to the data sources (Chiang and Zhang, 2016). It can consist of multiple hierarchical edge clusters, ranging from resource-constrained devices to micro data centers (Jansen et al., 2023), and networking entities, such as access points, gateways, and routers (Antonini et al., 2019). Containing diverse and heterogeneous devices with varying capabilities, fog creates new opportunities for system design to achieve better trade-offs between local and global data processing as well as distributed and centralized architectures (Chiang and Zhang, 2016).

Chiang and Zhang (Chiang and Zhang, 2016) state that the most fundamental question for fog computing is on where, when, and how to distribute system functions along the continuum. Closer proximity to data sources and end devices helps to meet the requirements of low latency, reduced bandwidth, and privacy guarantees enhancing the overall system efficiency and performance (Antonini et al., 2019). The authors of (Chiang and Zhang, 2016) also discuss the security tradeoffs of fog architectures. On the one hand, fog nodes can improve security by providing security functions, such as access control or encryption, for privacy-sensitive data before it leaves the edge. On the other hand, the operating environment is more vulnerable and resources to protect the system are more limited compared to the cloud.

Edge consists of decentralized, heterogeneous, and possibly mobile devices that typically have limited compute and storage resources (Jansen et al., 2023). The deployment of systems on the edge plays a key role in IoT solutions, and enable building complex cost-efficient applications with low latency (Antonini et al., 2019). Additionally, edge deployments can work reliably without an Internet connection (Alkhabbas et al., 2020) and preserve privacy (Jansen et al., 2023) making them suitable especially for autonomous vehicles and healthcare.

Although the edge addresses challenges inherent to the cloud, such as bandwidth limitations, increased latency, and high energy consumption, it introduces a new set of challenges that emerge during system development and maintenance. Developing and maintaining edge systems requires specialized expertise and skills from engineers (Alkhabbas et al., 2020). Moreover, the edge environment lacks standardized guidelines and foundational infrastructure services typical to the cloud (Jansen et al., 2023).

Table 2: The benefits and challenges of CC layers summarized.

	Benefits	Challenges
Cloud	Security and reliability Scalability of resources High processing power Standardized environment	Always-on connection required Geographical distance Delay and latency High transaction costs High energy consumption
Fog	Tradeoff between local and global Centralization/decentralization Privacy guarantees Varying capabilities	Heterogeneity of infrastructure May require technical skills High initial costs
Edge	Privacy guarantees Low latency Energy efficient Enables distribution Autonomy and control	High initial costs Constrained resources Vulnerable operation environment Heterogeneity of infrastructure Requires technical skills

3.2 Requirements of Data Platform Building Blocks

Data Spaces Support Centre (Data Spaces Support Center, 2024) provides a blueprint for data spaces including building blocks that divide platform capabilities into manageable clusters. We utilized the technical building blocks to evaluate the architectural fit between FDPs and CC. The building blocks contain the core functionality of FDPs.

- **Data models** enables the publishing, browsing, managing, and storing of shared dictionaries to facilitate semantic interoperability.
- **Data exchange** establishes mechanisms for the transmission of data according to the agreed protocol.
- **Provenance and traceability** provide capabilities for tracking data, its sharing, and usage for compliance purposes.
- **Identity and attestation management** enables platform actors to present, verify, store, and exchange attestations in a secure and self-sovereign manner.
- **Trust framework** provides criteria for platform actors, processes, and technical means for their validation and verification, and accredited trust sources.
- **Access and usage policies enforcement** establishes mechanisms to govern data management by implementing and enforcing policies for data access and usage.
- **Data, services and offering descriptions** provides tools for the creation, evaluation and maintenance of high-quality metadata to facilitate discovery, interoperability and usability.

- **Publication and discovery** enables dynamic transactions that bring together data/service providers and potential customers.
- **Value creation services** provide guidelines to define the technical infrastructure of services and set up technical conditions to ensure all value creation services have proper management, performance, scalability, monitoring and maintenance.

We derived relevant requirements from the building block descriptions to shape architectural decisions. Table 3 highlights the most relevant high level requirements for each building block. The highlighted requirements do not constitute an exhaustive list and should be interpreted as indicative rather than comprehensive. Nevertheless, they emphasize the most essential aspects relevant to placement decisions for each building block. In the other words, all listed requirements can be relevant for a building block, but the idea is to highlight only the ones that drive the placement decision. The requirements are the following.

- **Scalability** of storage and compute capacity ensure the extendability of a service and processing of high data volumes.
- **Consistency and manageability** refer to the ability of a service to maintain data synchronized across the platform.
- **Security and privacy** to protect actors and their business secrets.
- **Reliability** refers to the availability and fault resilience of a service that is closely related to business operations.
- **Low latency** enables real-time applications and reflects minimal cumulative transaction delays.
- **Low transaction costs** keep services with continuous transactions and high data volumes affordable.

- **Adaptability** refers to the ability of a service to accommodate the diverse technologies employed by different system components.
- **Performance** refers to the requirement for efficient operation without the need for particularly low latency, as the service is not used continuously.
- **Sovereignty** emphasizes an individual’s autonomy and ability to retain control over its data.

4 FEDERATED DATA PLATFORMS ON THE CLOUD CONTINUUM

By analyzing and comparing the requirements of FDP building blocks to the characteristics of CC layers, we were able to provide architectural guidelines for the service placement problem. Figure 2 summarizes the findings and indicates the optimal placements for each building block along the CC.

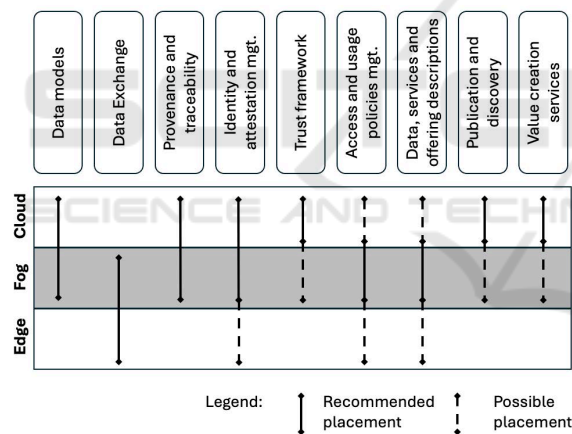


Figure 2: Building block placements.

We recommend that Data models building block be placed on the cloud or the higher levels of fog. All entities should be able to access and refer to the same set of data models, and thus a centralized service is preferred. Centralization will also ease the management of models, reducing duplicates and overlap. The number of data models to be stored will depend on the application, but it can be assumed that the number will remain bounded in every case. The functions related to this building block are not time-critical or require continuous transactions. Thus, deploying the service at the edge is not justified and may increase the complexity and cumulative costs of the system, as well as lead to divergence of models.

Data exchange building block benefits from deployment near the edge, as it provides functions enabling and supporting the actual exchange and sharing of data. Routing data through a centralized service can help to address interoperability challenges and provide a single point of access to all available data on the platform. However, a centralized cloud service is not suitable for cases involving low latency requirements, high data volumes, or sensitive data. Fog, on the other hand, can host such services by reducing high transaction costs and improving the overall performance. However, a centralized service creates a single point of failure that poses a significant risk to the operation of the system. Decentralized deployment approach at the edge or near edge can help minimize such risks and reduce cumulative latency implications as well as energy consumption. However, it increases the complexity of the system due to heterogeneous transfer protocols and adds local administrative burden regarding access and usage control.

Provenance and traceability functions are recommended to be deployed in the cloud or fog. A centralized cloud service or a predominantly centralized fog service can provide global transparency while guaranteeing the scalability and manageability of the platform. Global view and control over the devices involved ease the tracking. However, storing records related to the processes in remote systems can cause increased privacy risks as they may reveal business critical or otherwise sensitive information. The edge can offer a secure and privacy-preserving solution, but otherwise we do not see any significant benefits in that option.

Identity and attestation management can benefit from a distributed approach across the CC. Identity and attestation management building block ensures the integrity and reliability of information related to participating entities and their identity on the platform (Data Spaces Support Center, 2024). Such services are usually implemented centrally and benefit from the existence of a central registry. This building block does not include time-critical functions or large amounts of data transfers. However, identity data requires secure storage and transfer which are more feasible near the edge. Moreover, a service deployed closer to the edge in a more decentralized way can emphasize the principle of federation by which entities maintain strong autonomy and control over their data.

We recommend Trust framework be deployed in the cloud. The functions of the building block do not have strict latency requirements, not store highly sensitive data, and not require continuous transactions with other services. This indicates that deployment at

Table 3: The requirements of technical building blocks.

	Scalability	Consistency and manageability	Security and privacy	Reliability	Low latency	Low transaction costs	Adaptability	Performance	Sovereignty
Data models	x	x						x	
Data exchange			x	x	x	x	x		
Provenance and traceability	x		x					x	
Identity and attestation management	x	x	x					x	x
Trust framework	x	x						x	
Access and usage control enforcement			x	x	x		x		
Data, services, and offering descriptions	x		x					x	x
Publication and discovery	x	x	x					x	
Value creation services	x		x	x				x	

the edge does not bring any significant benefits, but can introduce unwanted complexity to the system. A cloud-based solution, on the other hand, can provide better control and transparency and ensure uniform criteria as well as scalability. Fog layers closer to the cloud offer similar benefits as the cloud, making them a viable alternative.

Access and usage policy enforcement includes functionalities that can benefit both, cloud and edge, thus making fog an optimal option. This building block provides functionality to specify and manage policies and track compliance (Data Spaces Support Center, 2024). There are various types of policies: some apply across the platform, while others are highly specific to data products. On one hand, policies operationalize legal data sharing contracts between involved parties, and therefore it is logical for the functionality to reside at the same level where such contracts are established. On the other hand, compliance tracking is tightly related to data transfer and usage processes that likely take place at the edge. However, deploying the full functionality at the edge can add significant administrative load. Utilizing fog enables a semi-centralized approach that reduces administrative burden while keeping the control close to the entities.

Fog serves as a good compromise for Data, services, and offering description building block. This building block offers tools for metadata management, validation and version control (Data Spaces Support Center, 2024). Metadata management drives the service placement decisions closer to the edge, whereas validation and version control functionalities can ben-

efit more from deployment on fog or cloud levels. The edge offers full visibility to the actual data enabling quick response to changes. Moreover, metadata management close to the corresponding entity reinforces its autonomy and control. In contrast, a cloud-based solution could provide secure and reliable storage capabilities for version control and ensure the interoperability of descriptions through a centralized validation service. Considering this, fog can serve as a good compromise if the functionalities of the building block are intended to be kept centralized rather than distributed along the CC.

Publication and discovery building block benefits most from a centralized cloud-based solution. From the perspective of system management and scalability, a centralized cloud deployed meets the requirements in many cases. It facilitates the discovery of available offerings and ensures scalability for large amounts of offerings. Moreover, this building block does not contain any time-critical functions or require continuous transactions that would drive placement decision closer to the edge. Privacy, however, is an aspect that could justify deploying the service on the fog. Distributing the functionality closer to the edge can create unnecessary processing that will reduce the overall performance of the system. It can also limit the ability of entities to establish new provider-customer relationships on the platform.

There is a strong tendency for value creation services to align with and utilize cloud capabilities. This building block provides support for the definition, provision, delivery, and utilization of services, which aim to create value out of data, such as visualization

and analytics, data quality services, additional security functions, and monitoring (Data Spaces Support Center, 2024). The support services has to be easily accessible, and they do not create continuous transactions or require low latencies. Thus, a centralized service deployed in the cloud or fog is optimal. The actual value creation services themselves may have specific requirements. Based on our experience, many of these kinds of supporting services benefit from the capabilities offered by the cloud. However, the specific requirements of each case must be carefully considered when deployment environments are evaluated.

5 APPLICATION IN MANUFACTURING AND INTRALOGISTICS

First, manufacturing and intralogistics systems need to serve the manufacturing process by monitoring and controlling the production and asset flow in real time. This is the primary functional requirement for service placement. Using fog and edge computing can increase quality of service (less delay, higher task completion ratio), improve efficiency (less network congestion, lower costs in data transfer and cloud computing), and improve process and personnel security as data is not transferred to cloud for analysis and decision making (Jansen et al., 2023) (Natesha and Gudeti, 2021).

Second, mobility likely plays a smaller role in performance of the FDP than in many other cloud computing deployments. Reversing the logic of (Ramírez et al., 2017), the movement and movement patterns happen in a limited and controlled production environment, and the end user data and service requests can be standardized which can reduce system volatility considerably. The controlled nature of the operating environment does not impose significant constraints or challenges on the use of cloud services.

Third, power consumption is another characteristic affecting the performance of CC deployments (Ramírez et al., 2017). As manufacturing and intralogistics systems are connected to physical manufacturing and transport processes, the power consumption of ICT systems can be considered to be marginal compared to those of actual production processes, when the alternative costs could be poor production performance or quality.

Fourth, manufacturing and intralogistic applications can complicate deployment with their large scale and heterogeneity of IoT devices, sensors and actuators (Salaht et al., 2020). Geographical distri-

bution does not come into play with manufacturing and intralogistics scenarios as one production facility typically can be considered to operate in their own fog layer, access cloud services only when needed. A single fog service placement problem is considerable simpler (Ramírez et al., 2017).

Finally, even though manufacturing and intralogistics scenarios can be simple in some aspects, it is still challenging to determine where and when to deploy each of the computing tasks to meet all functional and non-functional requirements and to optimize for several interdependent parameters (Brogi et al., 2020). Therefore, more detailed application-specific requirements must be identified through additional research methods before defining service placements.

6 DISCUSSION

The results show that federated data platforms benefit from the CC. Most of the building blocks represent functionalities for which a centralized cloud deployment is preferred due to reduced system complexity and increased scalability. However, we identified two cases where key functionalities can benefit significantly from edge deployment.

Firstly, some building blocks require stronger privacy protection. In such cases, centralized services can be hosted at the fog layer, preventing highly sensitive data from being transferred to remote systems. Furthermore, hosting the services closer to the edge facilitates data sovereignty which is one of the key design principles for FDPs discussed by Jarke et al. (2019). However, the placement decision comes with trade-offs, as the responsibility for security metrics shifts from a cloud provider to IoT engineers, as noted by Alkhabbas et al. (2020). While industrial application environments are typically well-secured, this shift demands additional skills and expertise from organizations.

Secondly, the only building block benefiting from a fully distributed deployment at the edge is Data exchange, especially in applications involving high volumes of data, such as manufacturing and intralogistics. Processing high volumes of data in the cloud can compromise real-time decision making and reliable operation that are essential for industrial applications (Ahearne et al., 2023) (Jansen et al., 2023). Beyond performance considerations, the costs of data transfer also justifies deploying the service at the edge, as highlighted by Alkhabbas et al. (2020).

The presented analysis framework and the results help practitioners evaluate different service placement

options, thus facilitating the design of future federated data platforms. The framework can also be utilized when assessing the architectural designs of existing systems. The proposed service placements and the analysis provide a clearer understanding of the benefits of different design choices and emphasize the advantages of service deployment across the CC.

The methodological choices impose limitations on this study. Firstly, this study is only based on a theory and practical implications of proposed service placements have not been tested. This can affect the reliability of the results in practice. Secondly, the levels of the CC are not as clearly defined in reality as the proposed framework suggests. Modern edge nodes can be capable in terms of storage capacity and processing power, enabling them to host services with substantial resource requirements. Lastly, the considered requirements of each platform building block are based on a short general description and the assumptions of its functionality. These assumptions reduce the accuracy of the proposed service placement recommendations. To enhance the accuracy, both application specific requirements and the capabilities of the available hardware should be taken into consideration.

This study is the first step towards evaluation of the benefits, challenges and applicability of placement of FDPs along the CC. While we see the logical fit of FDP components with the characteristics of cloud continuum layers, we do not take into account the restrictions of underlying infrastructure nodes and communication network. Future research should focus on validating the recommended service placements by applying the framework across a variety of industrial use cases. Alternatively, these placements and their implications can be further analyzed with simulations and algorithmic analysis. Secondly, our analysis is not yet addressing the dynamic allocation and reallocation of compute tasks which is a core benefit of cloud continuum. Defining a method and algorithms for dynamic placement of FDP building blocks and scheduling tasks is an unexplored research area. Further studies can evaluate different allocation strategies by measuring metrics, such as latency, energy efficiency, and cost, across various workloads. Furthermore, the impacts of data placement options can be further explored, i.e. how deliberate placement of data in edge or fog nodes impacts the system performance.

7 CONCLUSION

Data sharing between different systems and stakeholders enables significant economic benefits to or-

ganizations. Federated data platforms facilitate the sharing, integration, and management of data between diverse data sources and stakeholders. However, the deployment of FDPs in a centralized cloud introduces challenges in application areas, such as manufacturing, that produce a large volume of data in resource-constrained environments with low latency requirements. To address this challenge, we studied architectural fit between FDPs and the CC to deploy platform services closer to data sources and users.

We designed and applied an analysis framework for evaluating the architectural fit. The framework is based on the characteristics of different CC levels and the identified requirements of FDP building blocks described by Data Spaces Support Centre. The results show that deploying services directly connected to high volumes of data closer to the edge can improve overall performance and efficiency of a system. In contrast, some of the services are centralized by their nature and the similar placement decision could increase complexity making the system hard to manage.

ACKNOWLEDGEMENTS

This research has been funded by Business Finland research project Twinflow.

REFERENCES

- Ahearne, S., Khalid, A., Ron, M., and Burget, P. (2023). An ai factory digital twin deployed within a high performance edge architecture. In *2023 IEEE 31st International Conference on Network Protocols (ICNP)*, pages 1–6. IEEE.
- Al-Dulaimy, A., Jansen, M., Johansson, B., Trivedi, A., Iosup, A., Ashjaei, M., Galletta, A., Kimovski, D., Prodan, R., Tserpes, K., et al. (2024). The computing continuum: From iot to the cloud. *Internet of Things*, 27:101272.
- Alkhabbas, F., Spalazzese, R., Cerioli, M., Leotta, M., and Reggio, G. (2020). On the deployment of iot systems: An industrial survey. In *2020 IEEE International Conference on Software Architecture Companion (ICSA-C)*, pages 17–24. IEEE.
- Ammann, J. and Hess, T. (2025). To sell, to donate, or to barter? value creation and capture through business model types in decentralized data ecosystems. *Electronic Markets*, 35(1):1–22.
- Antonini, M., Vecchio, M., and Antonelli, F. (2019). Fog computing architectures: A reference for practitioners. *IEEE Internet of Things Magazine*, 2(3):19–25.
- Barbosa, V., Sabino, A., Lima, L. N., Brito, C., Feitosa, L., Pereira, P., Maciel, P., Nguyen, T. A., and Silva, F. A.

- (2025). Performance evaluation of iot-based industrial automation using edge, fog, and cloud architectures. *Journal of Network and Systems Management*, 33(1):1–25.
- Bellavista, P., Foschini, L., Ghiselli, N., and Reale, A. (2019). Mqtt-based middleware for container support in fog computing environments. In *2019 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–7. IEEE.
- Bonomi, F., Milito, R., Natarajan, P., and Zhu, J. (2014). Fog computing: A platform for internet of things and analytics. *Big data and internet of things: A roadmap for smart environments*, pages 169–186.
- Brogì, A., Forti, S., Guerrero, C., and Lera, I. (2020). How to place your apps in the fog: State of the art and open challenges. *Software: Practice and Experience*, 50(5):719–740.
- Byers, C. and Swanson, R. (2017). Openfog consortium openfog reference architecture for fog computing. *OpenFog Consortium Archit. Working Group, Fremont, CA, USA, Tech. Rep. OPFRA001*.
- Cardellini, V., Dazzi, P., Mencagli, G., Nardelli, M., and Torquati, M. (2025). Scalable compute continuum.
- Chiang, M. and Zhang, T. (2016). Fog and iot: An overview of research opportunities. *IEEE Internet of things journal*, 3(6):854–864.
- Čilić, I., Žarko, I. P., and Kušek, M. (2021). Towards service orchestration for the cloud-to-thing continuum. In *2021 6th International Conference on Smart and Sustainable Technologies (SpliTech)*, pages 01–07. IEEE.
- Curry, E., Scerri, S., and Tuikka, T. (2022). *Data spaces: design, deployment and future directions*. Springer Nature.
- Data Spaces Support Center (2024). Data spaces blueprint v2.0 - home. <https://dscc.eu/space/BVE2/1071251457/Data+Spaces+Blueprint+v2.0++Home>, Last accessed on 2025-04-30.
- Di Martino, B., Pezzullo, G. J., and Beggato, C. (2023). Towards optimized design and deployment of a military supply chain on federated cloud continuum supported by simulation-based performance evaluation. In *2023 IEEE International Workshop on Technologies for Defense and Security (TechDefense)*, pages 423–428. IEEE.
- Du, X., Tang, S., Lu, Z., Wet, J., Gai, K., and Hung, P. C. (2020). A novel data placement strategy for data-sharing scientific workflows in heterogeneous edge-cloud computing environments. In *2020 IEEE International Conference on Web Services (ICWS)*, pages 498–507. IEEE.
- European Commission (2020). A european strategy for data: communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions. <https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy-en>, Last accessed on 2024-11-30.
- European Commission (2024). Common european data spaces. <https://digital-strategy.ec.europa.eu/en/policies/data-spaces>, Last accessed on 2024-11-24.
- Farhadi, V., Mehmeti, F., He, T., La Porta, T. F., Khamfroush, H., Wang, S., Chan, K. S., and Poularakis, K. (2021). Service placement and request scheduling for data-intensive applications in edge clouds. *IEEE/ACM Transactions on Networking*, 29(2):779–792.
- Gelhaar, J., Groß, T., and Otto, B. (2021). A taxonomy for data ecosystems. In *Hawaii International Conference on System Sciences (HICSS) 2021*.
- Gilbert, S. and Lynch, N. (2002). Brewer’s conjecture and the feasibility of consistent, available, partition-tolerant web services. *Acm Sigact News*, 33(2):51–59.
- Jansen, M., Al-Dulaimy, A., Papadopoulos, A. V., Trivedi, A., and Iosup, A. (2023). The spec-rg reference architecture for the compute continuum. In *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, pages 469–484. IEEE.
- Jarke, M., Otto, B., and Ram, S. (2019). Data sovereignty and data space ecosystems.
- Jonas, E., Schleier-Smith, J., Sreekanti, V., Tsai, C.-C., Khandelwal, A., Pu, Q., Shankar, V., Carreira, J., Krauth, K., Yadwadkar, N., et al. (2019). Cloud programming simplified: A berkeley view on serverless computing. *arXiv preprint arXiv:1902.03383*.
- Jussen, I., Fassnacht, M., Schweihoff, J. C., and Möller, F. (2024). Reaching for the stars: Exploring value constellations in inter-organizational data sharing. *ECIS 2024 Proceedings*. 3.
- Kasi, S. K., Kasi, M. K., Ali, K., Raza, M., Afzal, H., Lasebae, A., Naeem, B., Ul Islam, S., and Rodrigues, J. J. (2020). Heuristic edge server placement in industrial internet of things and cellular networks. *IEEE Internet of Things Journal*, 8(13):10308–10317.
- Kimovski, D., Mathá, R., Hammer, J., Mehran, N., Hellwagner, H., and Prodan, R. (2021). Cloud, fog, or edge: Where to compute? *IEEE Internet Computing*, 25(4):30–36.
- Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., and Hoffmann, M. (2014). Industry 4.0. *Business & information systems engineering*, 6:239–242.
- Masip-Bruin, X., Marín-Tordera, E., Tashakor, G., Jukan, A., and Ren, G.-J. (2016). Foggy clouds and cloudy fogs: a real need for coordinated management of fog-to-cloud computing systems. *IEEE Wireless Communications*, 23(5):120–128.
- Moreschini, S., Pecorelli, F., Li, X., Naz, S., Hästbacka, D., and Taibi, D. (2022). Cloud continuum: The definition. *IEEE Access*, 10:131876–131886.
- Natesha, B. and Guddeti, R. M. R. (2021). Adopting elitism-based genetic algorithm for minimizing multi-objective problems of iot service placement in fog computing environment. *Journal of Network and Computer Applications*, 178:102972.
- Otto, B. (2022). A federated infrastructure for european data spaces. *Communications of the ACM*, 65(4):44–45.

- Otto, B., Jürjens, J., Schon, J., Auer, S., Menz, N., Wenzel, S., and Cirullies, J. (2016). Industrial data space. digital sovereignty over data.
- Otto, B., Ten Hompel, M., and Wrobel, S. (2022). *Designing data spaces: The ecosystem approach to competitive advantage*. Springer Nature.
- Pandey, A. K. and Pandey, R. (2020). Influence of cap theorem on big data analysis. *Int. J. Inform. Technol. (IJIT)*, 6(6).
- Poularakis, K., Llorca, J., Tulino, A. M., Taylor, I., and Tassioulas, L. (2020). Service placement and request routing in mec networks with storage, computation, and communication constraints. *IEEE/ACM Transactions on Networking*, 28(3):1047–1060.
- Rajput, S. and Singh, S. P. (2019). Connecting circular economy and industry 4.0. *International Journal of Information Management*, 49:98–113.
- Ramírez, W., Masip-Bruin, X., Marin-Tordera, E., Souza, V. B. C., Jukan, A., Ren, G.-J., and de Dios, O. G. (2017). Evaluating the benefits of combined and continuous fog-to-cloud architectures. *Computer Communications*, 113:43–52.
- Richter, H. and Slowinski, P. R. (2019). The data sharing economy: on the emergence of new intermediaries. *IIC-International Review of Intellectual Property and Competition Law*, 50:4–29.
- Salaht, F. A., Desprez, F., and Lebre, A. (2020). An overview of service placement problem in fog and edge computing. *ACM Computing Surveys (CSUR)*, 53(3):1–35.
- Santos, J., Wauters, T., Volckaert, B., and De Turck, F. (2021). Towards low-latency service delivery in a continuum of virtual resources: State-of-the-art and research directions. *IEEE Communications Surveys & Tutorials*, 23(4):2557–2589.
- Shen, Z., Liu, B., and Dou, W. (2024). An effective data placement strategy for iiot applications. *Concurrency and Computation: Practice and Experience*, 36(10):e7977.
- Sousa, R., Nogueira, L., Rodrigues, F., and Pinho, L. M. (2022). Global resource management in the elastic architecture. In *2022 IEEE 5th International Conference on Industrial Cyber-Physical Systems (ICPS)*, pages 01–06. IEEE.
- Suleiman, Z., Shaikholla, S., Dikhanbayeva, D., Shehab, E., and Turkyilmaz, A. (2022). Industry 4.0: Clustering of concepts and characteristics. *Cogent Engineering*, 9(1):2034264.
- Trakadas, P., Masip-Bruin, X., Facca, F. M., Spantideas, S. T., Giannopoulos, A. E., Kapsalis, N. C., Martins, R., Bosani, E., Ramon, J., Prats, R. G., et al. (2022). A reference architecture for cloud–edge meta-operating systems enabling cross-domain, data-intensive, ml-assisted applications: Architectural overview and key concepts. *Sensors*, 22(22):9003.
- Xiao, Z., Song, W., and Chen, Q. (2012). Dynamic resource allocation using virtual machines for cloud computing environment. *IEEE transactions on parallel and distributed systems*, 24(6):1107–1117.
- Yousefpour, A., Fung, C., Nguyen, T., Kadiyala, K., Jalali, F., Niakanlahiji, A., Kong, J., and Jue, J. P. (2019). All one needs to know about fog computing and related edge computing paradigms: A complete survey. *Journal of Systems Architecture*, 98:289–330.