Driving Industry 4.0 at Distributed Edges with Cloud Orchestration

Authors:

Thiago Weber Martins
Industry Standards
Open Source and Industry 4.0
Innovation Architect
SAP SE
thiago.weber.martins@sap.com

Dr. Frank Rambo
SAP IoT Product
Management
SAP SE
frank.rambo@sap.com

Christian Viezens
Program Lead - Edge
Lifecycle Management
SAP SE
christian.viezens@sap.com

Erich Clauer
VP of Industry Standards,
Open Source and Industry 4.0
Innovation
SAP SE
erich.clauer@sap.com

Andreas Sandner
Vice President - SAP DSC
Manufacturing and Industrial
IoT Engineering
SAP SE
andreas.sandner@sap.com

Tim McConnell
Development Architect -
SAP DSC Manufacturing and
Industrial IoT Engineering
SAP SE
tim.mcconnell@sap.com

Bakhtiyar Mirtov
Development Architect -
SAP DSC Manufacturing and
Industrial IoT Engineering
SAP SE
bakhtiyar.mirtov@sap.com

Robert Noce
Chief Development Architect,
Digital Supply Chain
SAP SE
robert.noce@sap.com

Stamatis Karnouskos
Industry Standards, Open
Source and Industry 4.0
Innovation
SAP SE
stamatis.karnouskos@sap.com

INTRODUCTION

Industry 4.0 is disrupting the industry and is changing the *linear fixed-purpose architectures towards decentralized processes that enable swarms of intelligent assets with autonomous capabilities (machine to machine communication and interaction) that can be leveraged dynamically within highly-networked and diverse environments*. This transformation can be attributed in part to the convergence of Operational Technology (OT) with Information Technology (IT) systems that is reinventing on how participants within the whole supply chain communicate and interact. While in the past their communication and interaction patterns were structured in a hierarchal manner e.g., following the ISA-95 automation pyramid (Figure 1, left), Industry 4.0 is making these patterns more flexible, networked, and decentralized (Figure 1, right).

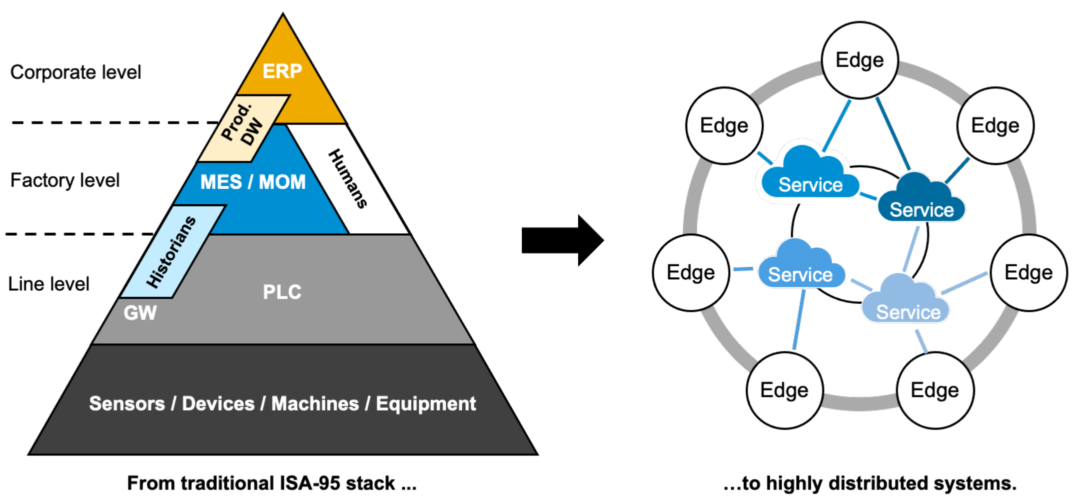

Fig. 1 - Industry 4.0 is driving a shift from a hierarchical and centralized stacks to highly distributed systems.

Besides considering the chances of highly networked and distributed systems, modern architectures for industrial setups must address complexity and be robust against disruptions. Disruptions in business-critical processes within the Design to Operate (D2O) phases (e.g., design, plan, procure, manufacture, deliver and operate) [1] may lead to costly interruptions of the whole supply chain. As the supply chain is highly networked around the globe, any interruption in a node can propagate to its next nodes in a domino effect.

Consider a manufacturing scenario with total loss of cloud connectivity that is coupled with last minute unplanned changes (shift worker unreliability, material delays, and resources unavailability). Thus, businesses need to also acquire, combine, and process data locally from multiple sources (OT with IT systems) in a disconnected mode and to execute business processes (in particular, mission critical and latency sensitive processes) independently of the cloud. At the same time, an architecture for Industry 4.0 must be able to support dynamic, data-driven production optimization.

Edge computing is a promising approach to mitigate risk and to avoid that production or assembly lines do not stop due to intermittent connection to the cloud. Typically, edge computing is performed in close proximity to the factory shop floor or warehouse, leveraging connectivity to both enterprise business solutions and onsite Manufacturing Execution Systems (MES), Programmable Logic Controllers (PLC), sensors, and other industrial automation systems. Consequently, this approach helps to achieve the targeted Key Performance Indicators (KPI) such as Overall Equipment Effectiveness (OEE). For these reasons, edge computing is an essential element in the reference architecture for Industry 4.0 [2]. However, it is crucial to develop strategies addressing the higher complexity in edge computing approaches and keeping the costs (e.g., development, integration, and operational) under control.

This paper introduces the approach SAP recommends to enable faster decision making on premises and to achieve continuity of business-critical processes in manufacturing by extending a subset of cloud capabilities to the edge, close to the source of data. Challenges related to deploying, managing, orchestrating, and operating distributed business applications in different manufacturing sites are considered and addressed to guarantee the targeted business outcome. This paper also discusses the approach and gives recommendations where to deploy business applications based on latency and criticality. Based on scenarios for Industry 4.0 introduced in [3], selected mission critical scenarios for logistics and manufacturing will be discussed to elaborate the "added value proposition" of the solutions presented in this paper. Furthermore, we give an overview to the possible changes and improvements on the shop floor and in the supply chain.

EDGE COMPUTING – CHARACTERISTICS, CHANCES AND CHALLENGES

Table 1 introduces the dimensions characterizing the edge computing and cloud computing paradigms. Their comparison emphasizes the chances and challenges to be addressed by applying edge computing approaches to manufacturing processes that will be discussed in this section.

Table 1 - Dimensions of edge computing and cloud computing.

Dimension	Edge	Cloud
Latency:	low	high
Number of nodes:	many	few
Computing and storage resources:	from constrained lightweight edge devices to heavy edge nodes with more resources	nodes with high to very high computing and storage resource (hyper-scaler)
Topology:	distributed	centralized
Complexity:	heterogeneous (devices, industry protocols)	homogenous

Potentials and Market Chances for Edge Computing Solutions

Edge computing is attracting a huge market potential due to its promising business value proposition. According to IDC Research [4] , by 2023 "*70% of IoT deployment will include edge-based decision making, supporting organizations' operational and strategic agendas and 70% of enterprises will run varying levels of data processing at the IoT edge. In tandem, organizations will spend over $16 billion on IoT edge infrastructure in that time*"*.* This estimation is not a surprise since the number of connected assets along the D2O phases growing.

A smart factory will be producing on average "*5 petabytes of data per day"* [5]. With 5G cellular networks, a higher volume of data produced by machines and robots can be available with higher throughput for real time scenarios in manufacturing such as smart sensing. In this context, such scenarios are highly dependent on "*human-like latency and always-on connectivity*" [6]. Analogically, the same capabilities are mandatory in autonomous transportation (e.g., autonomous driving). Both manufacturing and autonomous transportation scenarios will profit from artificial intelligence (AI) at large, and more specifically from machine learning (ML) and deep learning (DL) applications using edge computing capabilities to enable instantaneous decision making and optimize the processes in real time.

Furthermore, use cases and deployment scenarios that require real-time decision-making support within verticals (healthcare, smart cities, autonomous vehicles, smart retail) are augmented and virtual reality (AR/VR), Visual Inspection, Cloud Gaming [7]. In context of Industry 4.0, edge computing supports the realization of Intelligent Assets, Intelligent Factories, Intelligent Logistics, Intelligent Products, and Empowered People scenarios [3] into a holistic, intelligent supply chain. Further introduction to and description of Edge Computing for industrial applications can be found in the literature [8] [9] [10, 11].

Challenges of Edge Computing

Edge computing offers compelling benefits for Industry 4.0 scenarios. It also presents challenges that must be addressed.

From Few to Many Distributed Edge Nodes – Higher Complexity. Edge computing is characterized by highly distributed topologies that often consist, depending on the scenario, of edge nodes with limited computing and storage. At the same time, edge landscapes are highly heterogeneous in terms of platforms, infrastructure, hardware networking, and security [12]. Both characteristics increase the complexity of deployment, management, and orchestration of edge solutions. Heterogeneous landscapes raise some challenges such as interoperability and *plug and* play solutions to ensure the continuity of business-critical processes. To maximize business value, edge computing deployments need to be simplified to ensure that the edge focuses only on the business-critical and edge-specific components of business applications being extended from the cloud.

IT/OT convergence introduces design constraints to the edge. Typically, OT and IT solutions have different constraints and requirements that guide design and implementation. While OT focuses on safety, efficiency, and continuity of manufacturing processes, IT prioritizes speed, security, and analysis of data [13]. Modern edge computing must merge the qualities of both paradigms to maximize its potential. For example, how to leverage the agility of IT systems and ensure their high availability for the continuity mission critical and latency sensitive process at OT level. Additionally, challenges regarding backend data integration at the edge from different data sources (cloud, other edge nodes, and assets) and their semantics that are distributed within various systems on the shop-floor must be considered. These challenges raise the question on what data is business-critical and must be computed at the edge.

Integration of legacy systems (Brownfield). Existing equipment might not support greenfield scenarios using the newest information & communication technologies (ICT). Edge computing solutions must consider strategies such as retrofitting to support legacy scenarios and semantics.

Development, Integration and Operational Costs. Distributed systems without centralized management and orchestration strategies can have a high Total Cost of Ownership (TCO) impact. Development costs also need to be considered to enable custom development at the edge by providing the optimal balance between standardization, openness, and configurability of applications and services. Operational approaches need to be evaluated from a variety of perspectives. One perspective is the cost related to communications costs. Filtering and aggregation capabilities reducing the workload are decisive in this context.

In the next sections, we describe how we derive these challenges into requirements that drive to solutions for edge computing as elaborated in this paper.

Requirements for Edge Computing Solutions in Manufacturing

This section introduces requirements for *functional* and *non-functional* features of edge computing solutions addressing the challenges above.

Features fulfilling *functional requirements* to build edge computing capabilities are:

- Machine orchestration
- Connectivity of edge nodes to assets, other edge nodes, and cloud applications using industry standards (e.g., OPC-UA, MQTT, Modbus, REST, WebSockets)
- Edge analytics, data processing, and local storage
- Zero-touch device provisioning and onboarding
- Backend data integration and semantics (master and transactional data)
- Compliance to data security, privacy, sovereignty e.g., General Data Protection Regulation (GPDR)

Components fulfilling *non-functional requirements* to build an edge computing infrastructure and platform are:

- Offline capabilities (store and forward buffering, connection resiliency, etc.)
- Elasticity / Scalability
- Extensibility
- High availability and reliability
- Cloud orchestrated edge: lifecycle of distributed edge nodes based on its individual business configurations
- Automated, and continuous deployment, orchestration, management, replacement, recovery of edge nodes (CI/CD)

APPROACH AND ARCHITECTURE FOR BUSINESS APPLICATIONS ON DISTRIBUTED EDGES WITH CLOUD ORCHESTRATION

Addressing the requirements discussed above, this section introduces an effective approach and architecture for edge computing solutions enabling mission critical execution of distributed manufacturing processes based on [14] [15]. One of the main elements of this approach is moving business applications as well as business data (e.g., master data, manufacturing orders) closer to the manufacturing processes and combine them with OT data from multiple systems.

This approach aims at achieving faster decision-making closer where data is generated. Also, we elaborate on how an efficient edge computing architecture addresses the main limitations of cloud computing such as latency, limited bandwidth and intermittent connectivity that may result to costly disruptions within manufacturing processes. Finally, we introduce how continuous edge lifecycle management and centralized business configuration ensures low TCO as edge nodes and its workloads distributed in manufacturing processes are orchestrated from the cloud.

Edge-Cloud continuum for mission critical execution of distributed manufacturing processes. According to [12], the edge-cloud continuum is a physical infrastructure comprising of the internet, from discrete, decentralized devices to a centralized data center. Figure 2 shows an edge-cloud continuum focusing on enabling mission critical execution of distributed manufacturing processes on the edge. It considers the characteristics from Table 1 and is based on the following layers.

IoT devices layer consists of any device with sensing capabilities, connected with a network, ingesting data and to be controlled/orchestrated from the overlying layers.

Edge layer consists of a set of distributed physical edge devices (hardware such as gateways, hubs, PLCs, etc.), software components and the corresponding deployed business applications and data that forms an edge node. Physical edge devices are classified with different sizes (S, M, L, XL) depending on its computing and storage capabilities. Business applications are deployed on these devices and brings the necessary functionalities fulfilling a set mission critical and latency sensitive scenario. Example of applications are SAP Digital Manufacturing Cloud for edge computing [14] or AAS Service [15]. This layer is characterized (see dimensions of Figure 2) by fast response times (low latency) and context awareness based on ingested data from assets that is combined with business data from the cloud. Business applications running in this layer as well as its orchestration from the cloud are the focus of this paper and will be elaborated in the next subsections.

Cloud layer is the overlying computing infrastructure providing agile, flexible, on-demand access to business applications such as SAP Integrated Business Planning, Digital Manufacturing Cloud [16] [17], and Enterprise Resource Planning (ERP) systems. Cloud computing provides an affordable infrastructure for storing, managing, processing data in a centralized manner since its resources can be shared by multiple tenants. This layer is characterized (see dimensions of Figure 2) by higher latency and network dependency. The value added to data grows across this dimension since cloud is the ultimate centralized source for business data in cloud/edge solutions.

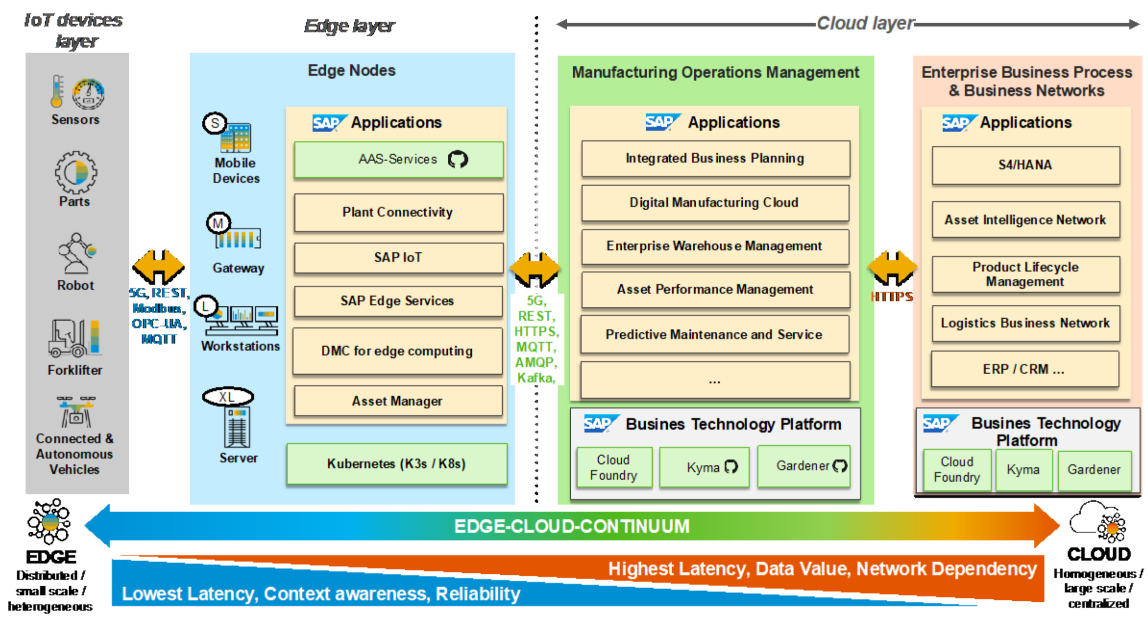

Fig. 2 - Edge-cloud continuum for enabling mission critical execution of distributed manufacturing processes

Approach for mission critical execution of distributed manufacturing processes on the edge. As illustrated in Figure 2, deployments of latency-sensitive and bandwidth-critical business applications are optimally placed in physical locations that promote the required continuity, robustness, and fast response times of business processes. The SAP Digital Manufacturing Cloud for edge computing approach [14] is an example of such deployment in the edge layer and is detailed below.

As shown in Figure 3, the approach consists, from an *infrastructure view,* of a cloud instance and one or more edge nodes that build the edge-cloud continuum as shown in Figure 2. Plants represent both the edge layer as well as its underlying assets. Many manufacturing companies run multiple production plants in different locations. Each plant needs to run an edge infrastructure within its own premise. Applications need to be deployed into each plant turning it into a distributed system. More granular scenarios can be achieved by deploying more than one edge infrastructure into a plant. For example, by deploying an edge infrastructure into each production line and all work centers of a single plant [18]. An edge infrastructure building an edge node consists of:

- *hardware* needed to provide the necessary computing resources (local processing and local storage),
- *software components* such as operating system, hypervisor, container runtime, and containerized business applications and backing services, and
- *interfaces* supporting industry standards (e.g., MQTT, OPC-UA, TCP/IP, REST) that establish bi-directional communication and interactions between cloud-edge layer, between edge nodes within the edge layer (east/west communication) as well as between assets-edge layer.

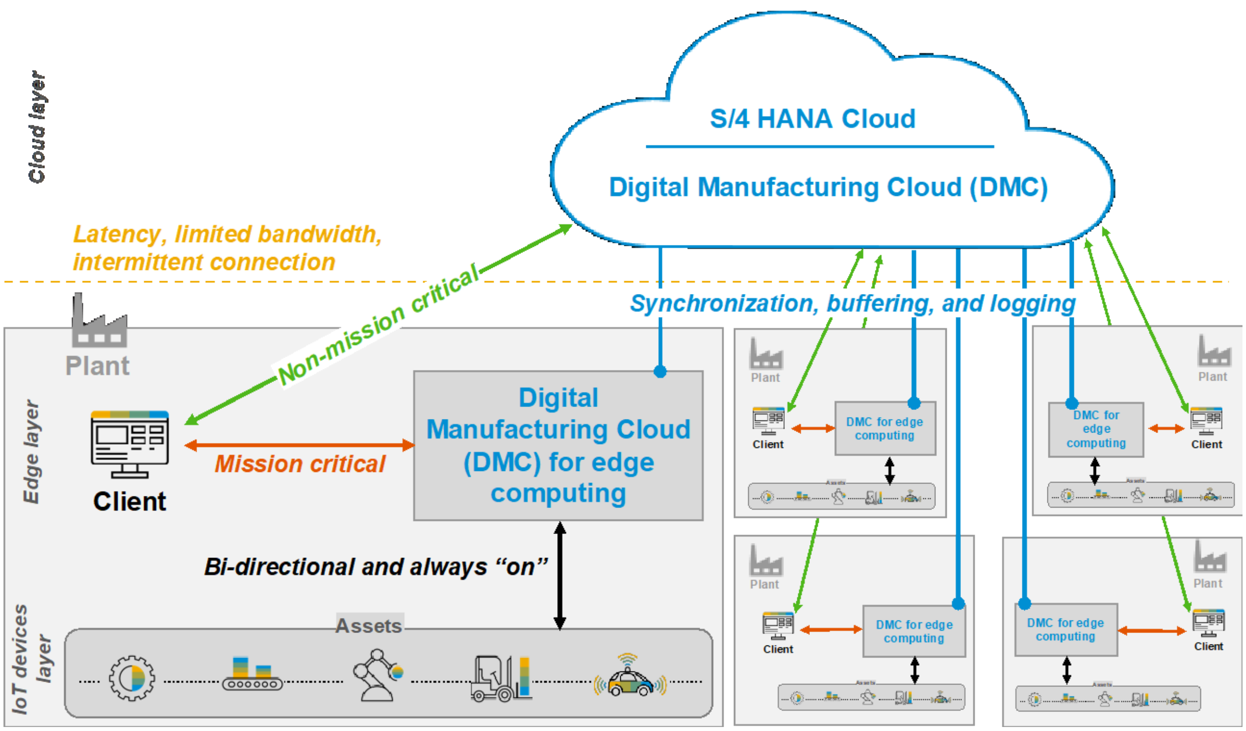

Fig. 3 - Approach for ensuring business continuity for mission critical applications based on [14].

As both cloud and edge provide cloud native capabilities such as support to container runtimes and containerized applications, the same business applications can run on the cloud or on the edge nodes in a distributed manner. This empowers solution providers with the flexibility to determine the optimal balance of cloud and edge functionality needed to fulfill their scenarios. The deployment of the containerized applications can vary by customer, depending on their infrastructure (e.g., edge hardware computing capabilities and footprint constraints) and needs. This enables OT/IT convergence, as the edge node with its deployed business applications is capable of ingesting raw data directly from the assets (OT data) and combining it locally with business context (IT data) made available from the cloud.

From a *usage and business scenario perspective,* this approach supports *manufacturing supervisors* and *operators* on the plant or shop floor with the ability to make faster decisions on manufacturing. For that, *non-mission critical applications* run in the cloud and *non-mission critical data* is stored in the cloud. If these functionalities and data are business-critical as well as latency-sensitive, they are additionally replicated to the corresponding edge nodes using local proxy / cache mechanisms. *Mission critical applications* can then run at the edge and *latency-sensitive data* are buffered on the edge for rapid retrieval.

Depending on the scenario and hardware constraints (e.g., footprint), these applications and/or this data are deployed on the corresponding edge nodes in a distributed manner. Edge nodes have then the capability to buffer, pre-process, and filter data before uploading it to, or downloading it from the cloud. It can optimize the data synchronization with cloud according to the current bandwidth, latency, and connection availability while also keeping connection costs down. It also ensures that any maintenance, update, or upgrade of the IT infrastructure does not affect the continuity of processes.

Relevant business applications for mission critical applications at the edge. After analyzing the manufacturing environment and the identification of the mission critical processes, it is necessary to define which business applications must run on the edge nodes. Digital Manufacturing Cloud (DMC) [16] is deployed in the cloud (Figure 3, top), while solutions for edge computing (DMC for edge computing [14] and Edge Services [19]) are deployed on the edge nodes (Figure 3, bottom).

Business applications in the cloud focus on functionalities such as resource orchestration and management, design of production processes, manufacturing execution, manufacturing networks, monitoring, and analytics. It also provides a central control plane for edge lifecycle management and orchestration of business applications, business data, and the corresponding business configuration for the underlying edge nodes.

Other examples of applications that are typically deployed in the cloud are manufacturing insights or training of AI/ML models for manufacturing inspection from a large amount of data that has been sent from the plant. Business applications deployed on the edge are *order management*, *manufacturing execution, dispatching, and monitoring*, *resource orchestration*, *dashboards*, *machine integration and automation*, *alerts,* and *execution of ML models*.

In addition to the business applications on the edge, *synchronization services for business objects and data* ensure that edge applications run independently of cloud connection. These services enable bi-directional movement of business data to provide eventual consistency between edge and cloud data sources. To provide optimal consistency, this service buffers data on either side when connection is not available. Once available, buffered data is sent to its destination in a manner that avoids distributed conflicts between edge and cloud systems.

Relevant data for mission critical applications at the edge. It is important that the necessary business data required to run these business-critical and latency sensitive applications are always available. Decisions on *what data* and *how often (intervals) it is synchronized with the cloud* might differ from one scenario to another. In context of manufacturing processes, *business application configuration data*, *master data*, *manufacturing orders*, and *execution order* are examples of data to be buffered on the corresponding edge nodes from the cloud. In context of mission critical execution, *master data at the edge* might consists of *Bill-of-Material (BOM)*, *routings*, *operation activities*, *materials*, *resources*, and *work centers*.

Depending on the complexity and criticality of the processes, enough business data should be sent from the cloud to the edge so that business applications can run "offline" within a pre-set time interval (e.g., n numbers of shifts or batches) without connectivity. Other requirements regarding compliance to data security, privacy, and sovereignty (e.g., GPDR) affect the choice of what data must be exclusively at the edge, synchronized with the cloud, or exclusively in the cloud. A detailed discourse on these topics, however, require further discussions in future works.

Architecture – Leveraging the capabilities of the cloud to the edge. This section details the edge computing architecture for mission critical and latency sensitive scenarios based on [21] [15]. The corresponding edge computing architecture is designed based on standardized and industrialized components that facilitate *plug and play* and easy deployment. One of the main innovations of the architectural approach is the combination of capabilities from cloud native and on-premises solutions to provide standardized business applications and reliable data synchronization for the edge.

Modern edge computing approaches leverage cloud native principles but addresses its distinguishing characteristics such as constraints on computing and storage resources, heterogenous ingestion, intermittent connectivity, and minimal latency. Business applications and orchestration infrastructure are developed and realized following modern paradigms such as containerization and industry de-facto standards such as Kubernetes.

Some benefits of cloud native principles are elasticity (ability to automatically scale computing resource regardless of current load to ensure the availability and performance of services), resilience (ability to provide a service even if certain components built into these services are not working), performance, modularity and re-usability, resource efficiency, automation, and observability. The result is high availability of services and the guarantee of designed response times. All these characteristics are relevant for distributed edge computing deployments as they address the specific requirements of business-critical manufacturing processes.

At the same time, the edge computing approach presented in this paper provides flexibility, extensibility, and individualization regarding deployment and configuration of business application at the edge. This includes lifecycle management and operation of applications at decentralized edge nodes from the cloud and will be detailed in the next section. The balance between standardization, ease of deployment as well as individualization reduce cost and simplifies custom edge applications development. This balance counts as one of the innovations provided by the solutions that are elaborated below.

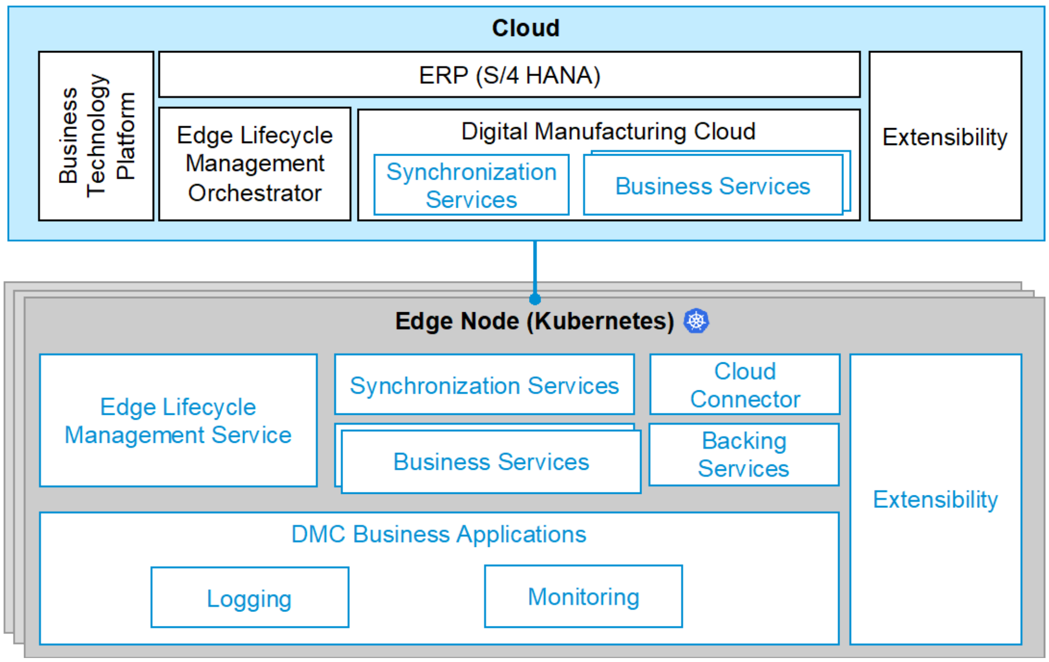

Fig. 4 - Edge computing architecture for mission critical execution of distributed manufacturing processes on the edge [20] [21].

Following these tenets, Figure 4 proposes an architecture and its components (business applications) based on [20] [21] that are allocated in the cloud and on the edge layer. Note that this architecture can be customized for specific needs and available resources.

The cloud layer contains business solutions, applications and services enabling centralized management and operation of the underlying edge nodes. The business applications and services run natively in containers on Kubernetes clusters. Modules of the architecture in this layer are:

- **S/4 HANA Cloud:** Intelligent Enterprise applications such as ERP system and database.
- **Business Technology Platform (BTP):** provides the runtime based on Cloud Foundry or Gardener [22] [23] for the business applications, APIs for interoperability and extensions, data models, process models for integration and workflow contents. [17]
- **Edge Lifecycle Management Orchestrator (ELMO):** standardized channel to provide lifecycle management of business applications and services from the cloud to the edge.
- **Digital Manufacturing Cloud (DMC):** manufacturing execution system (MES) deployed in the cloud providing a set of business applications and services to execute manufacturing

processes, analyze manufacturing and business data and integrate systems with low TCO. [16]

- **Extensibility:** supports execution relevant extensibilities scenarios via custom fields, API extensions, and functional extensions.

Edge nodes consist of local Kubernetes clusters running in the plant. Business applications and services that are essential for executing manufacturing processes are deployed on these clusters.

- **Cloud Connector:** establishes a secure connection between the cloud and the edge nodes to enable lifecycle management.
- **Edge Lifecycle Management Service:** counterpart of ELMO that is deployed on the edge to fulfill lifecycle tasks even if connection to the cloud is not available.
- **Synchronization Services:** provide services responsible for synchronizing data (configuration, master data, dynamic data (manufacturing order)) between edge node and cloud.
- **Business Services:** provide functionalities to enrich OT data streamed by assets with business context (data and processes) at the edge.
- **Backing Services:** provide services for data streaming and processing, data persistence, configuration management.
- **DMC Business Applications:** essential business applications from DMC that are deployed as containerized applications on the edge node and managed by edge lifecycle management services. Essential functionalities such as logging, and monitoring are also included.
- **Extensibility:** supports execution relevant extensibilities scenarios via custom fields, API extensions, and functional extensions at the edge.

High Availability for Edge Computing

Earlier implementations of edge computing solutions focused on lightweight data ingestion and processing scenarios. These workloads are handled in singular processes bound by the limitations of smaller resource footprints.

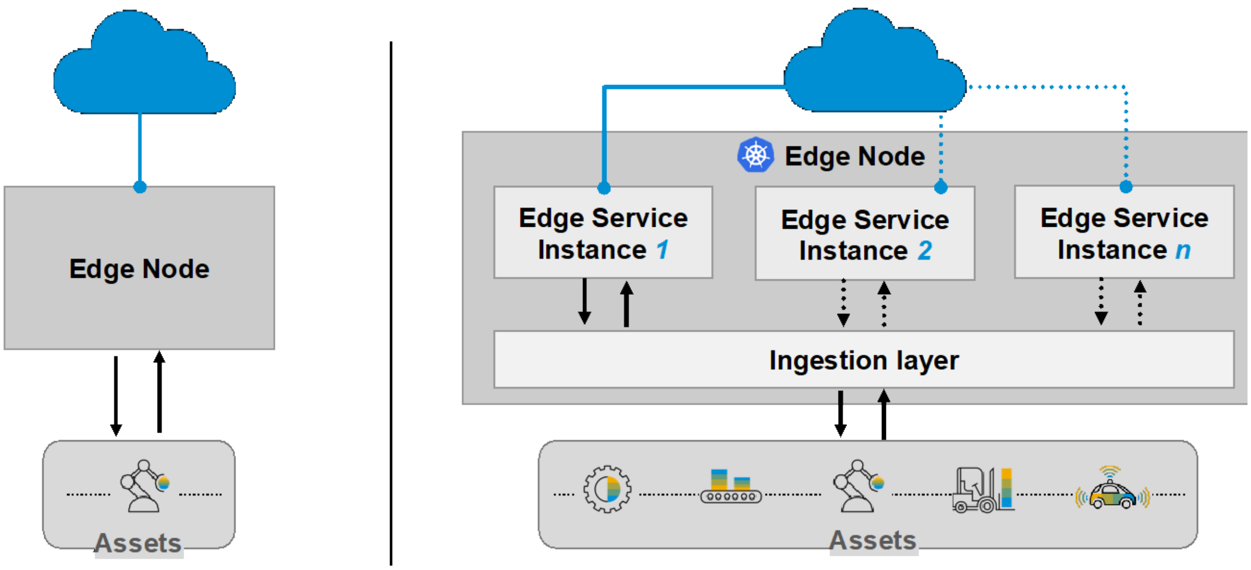

Fig. 5 - High Availability for Edge Computing (left: single edge node configuration; right: high-availability edge node configuration).

Transitioning to a Kubernetes-based platform allows to bring enterprise-grade applications to the edge in an easier and consistent manner. In addition, the existing edge computing modules are now evolved to be containers that are staged to be scaled horizontally to address increased utilization (see Figure 5). Additionally, Kubernetes provides a runtime that is open for setting up architecture's components such as backing services including databases and message brokers into High Availability (HA) configurations. This enables an edge architecture that can both scale elastically and provide stateful and stateless service redundancy to optimize response time and minimize downtime.

Edge Lifecycle Management Service and Orchestration

Often, the global nature of plant distribution can make managing and orchestrating edge nodes complicated and costly. Thus, it is necessary that modern edge computing architectures include functionalities to automate the management and orchestration of edge nodes within all steps of its lifecycle. In context of edge lifecycle management, it is important to keep business applications running at the edge always maintained and updated without needing to stop any process in the manufacturing (upgradeability).

As shown in Figure 4, Edge Lifecycle Management Service and Orchestration (ELMO) establishes standardized channel to provide lifecycle management of business applications and services from the cloud. Edge Lifecycle Management Service and Orchestration leverages de-facto standards such as Kubernetes and Helm.

A typical process for Edge Lifecycle Management is shown in Figure 6. The process can be structured in three lifecycle phases for an edge node: The onboarding phase, the operations phase and the decommission phase. During the *onboarding phase* Edge Lifecycle Management (EdgeLM) enables customers to register their newly received node from their hardware as a service provider in their customer tenant to make it operational.

In the *operations phase* EdgeLM provides central software lifecycle management capabilities including deploy, update, or remove containerized applications. It also enables monitoring of edge nodes and provides central logging capabilities. The customer factory manager can also locally perform emergency operations (start/stop) on the edge and can influence the scheduling when local updates are performed. When the *decommission phase* is reached, i.e., a customer decides to retire an edge node and cancel his hardware as a service subscription, EdgeLM offers capabilities to safely unregister the node from his cloud tenant.

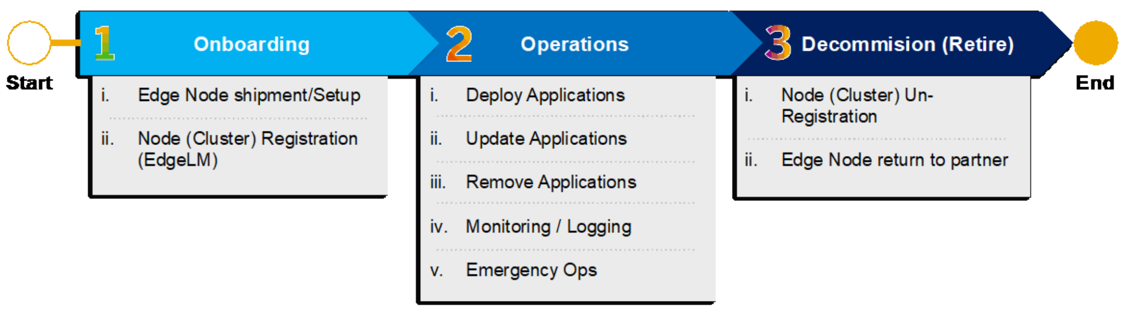

Fig. 6 - Phases of Edge Lifecycle Management.

Industry standards and open source driving the adoption of edge computing. As shown in the previous sections, many Industry 4.0 scenarios require interoperability and connectivity between assets and equipment on the shop floor as well as their interfaces to the overlaying edge-cloud infrastructure. Also, higher acceptance of new technologies as well as lower integration and operational costs are impacting the success of Industry 4.0 solutions in the market. Therefore, industry standards as well as open source driving de-facto standards are an imperative in modern edge computing architectures and are discussed in this section.

Examples of industry standards providing *interoperability and connectivity* for edge computing architectures are:

- **Asset - Edge Connectivity:** OPC-UA, MQTT, AMQP, DDS, REST, Modbus, WebSocket.
- **Edge - Cloud Connectivity:** REST, MQTT, AMQP, Apache Kafka, WebSocket.

Most of standards mentioned above are simultaneously being developed in standardization groups and realized within open source projects (page 86, Figure 27 of [24]). As discussed in [25], this trend leverages the potentials of open source, makes standardization more agile and transparent, and drives the adoption of new technologies. Examples are OASIS standards such as

AMQP and MQTT that are implemented in several open source projects from the Eclipse Foundation (Paho, Mosquitto), Apache (ActiveMQ) and, others like RabbitMQ and HiveMQ. Or OPC-UA implementations within the open source projects of OPC-UA Foundation [26] [27]. Eclipse IoT WG [28] drives many of the standard implementations as open source. Together with Docker containers, open source components such as Kubernetes and Helm are setting *industrial standard for building the infrastructure layers* of modern edge and cloud computing. Open source projects such as K3s or KubeEdge are adapting Kubernetes to the specific requirements of edge computing implementations. Other examples of open source initiatives driving de-facto standards for edge computing are: Eclipse Foundation's Edge Native WG [29] [30], Linux Foundation's LF Edge [31], and OpenInfra Foundation [32].

The *Asset-Administration Shell (AAS)* is a prominent industry standard defining *metamodels and semantics* (e.g., asset ID, attributes) to describe the asset it is representing (e.g., machines, AGVs, robots) and, hence, to realize the digital twin[1] in context of Industry 4.0 [36]. While it has initially been specified by the Plattform Industrie 4.0 [37] [38], the reference implementations are driven by the Industrial Digital Twin Association (IDTA) [39] [40] using open source as the development model [25]. The AAS consists of a set of submodels that allow the representation of specific aspects related to asset and its use case (e.g. technical data, operational data) [36] [41] [37] [38]. In general, AAS addresses the rather heterogenous edge computing landscape (devices, software, industry protocols).

Assets are then consistently described vertically in all layers (asset-edge-cloud) and horizontally in all phases from D2O. In this context, AAS facilitates and ensures the cross-company interoperability of on-boarding scenarios or the application, deployment, and orchestration of digital twin at the edge layer. Although AAS already provides many opportunities in terms of new scenarios and technologies for Industry 4.0 solutions, the specifications of these concepts are still evolving.

Thus, solutions exploring AAS at the edge to create autonomous, distributed systems still need to be discussed in future works. Particularly, how AAS supports backend data integration from different data sources (cloud, other edge nodes and assets distributed within various systems on the shop floor) using different standards (e.g., OPC-UA, AutomationML) at the edge.

In context of innovative data-driven business models for Industry 4.0, *International Data Spaces* (IDS) [42] [43] are setting the standards that allow a *sovereign and secure data exchange* between trusted partners in a federated approach [18]. At the plant (edge layer and IoT assets), IDS enables that data providers (owners) have a higher control over the data and its usage within industrial data spaces (e.g., Industry 4.0 data space [44]). IDS is being specified within the activities of

---

[1] Digital twin is defined as a digital representation of an entity, including its attributes and behaviors. The digital twin features bidirectional communication and interaction capabilities to its physical counterpart and environment, preferably in real-time, to ensure its as-manufactured or as-maintained representation. Further details on digital twins can be found in [51].

IDS-Association and being realized in several open source projects [45]. The question on how data sovereignty and the IDS concepts impact edge computing solutions goes beyond the scope of this paper and is subject of future contributions.

Initiatives such as Gaia-X [48] and Catena-X [49] are also considering edge computing as a fundamental infrastructure in their reference architecture. Gaia-X vision aims at establishing a federated data infrastructure fulfilling data sovereignty and security via standards and open source. Catena-X follows a similar vision of a secure and sovereign data space for companies and organizations involved in the automotive value chain. In both organizations, standards such as AAS and IDS are being considered. A detailed discourse on how these standards and initiatives such as Gaia-X and Catena-X help edge computing solutions to provide a better value to the industry should be topic of future works.

Detailed Logistics and Manufacturing Execution Scenarios

In this section, selected mission critical scenarios for logistics and manufacturing from [3] will be discussed using the edge computing approach introduce in the previous section. The goal is to elaborate the “added value proposition” of the solutions.

Smart Sensing at the Edge

Smart Sensing refers to automatically identifying physical objects through events, enriching them with business context and integrating them into business processes, without human intervention. Technologies typically considered as part of Smart Sensing include QR codes, RFID, optical character recognition, voice recognition and others.

Smart Sensing can identify movements of physical objects such as inbound or outbound deliveries on handling units such as pallets, containers, or tanks in real-time - even at high speed. When combined with edge computing, smart sensing can be leveraged to fully automate process steps in logistics and eliminate human errors. The edge nodes are key components as in real-life warehouse settings each scan needs to return immediate feedback to the warehouse worker. Turnaround time to the cloud interrupting the workflow on the ground can’t be tolerated.

Technically, cloud and edge components of SAP IoT work together to make this happen. Relevant hierarchical business context such as sales orders, deliveries, and handling units as well as customer and material master data is managed centrally in ERP – for example in S/4HANA. SAP IoT manages subscriptions to this business data and how smart sensing events are associated with individual handling units, for example. Decoupled synchronization between cloud and edge enables low-latency execution of smart sensing scenarios at the edge. Smart sensing events from handling units captured in a warehouse are ingested into the local edge node and set into its business context – now locally available at the edge. Business rules executed in the same edge node validate the events and provide instantaneous feedback to warehouse workers so that the work can continue. Asynchronously, the respective goods issue is posted to S/4HANA.

More advanced scenarios combine Smart Sensing with condition monitoring of sensitive goods leveraging IoT sensors, which ingest their readings into the same edge node used to process Smart Sensing events. Consider cold chains or delivery of fragile goods. Business rules monitor the continuous stream of IoT data from the sensor locally (at the edge) and only send IoT alerts to the cloud if beyond tolerances. These IoT alerts are seamlessly integrated into the user experience in S/4HANA. Rather than showing up in a siloed IoT app, the IoT alert appears, for example, embedded into the sales order app used by the targeted persona – for example a sales manager.

Edge Computing in Digital Manufacturing

The following scenario outlines the business value proposition of the edge computing solution presented above by applying it to solve mission critical manufacturing processes and to make supply chains more resilient. Figure 7 shows this scenario showcasing the manufacturing of an incoming order for a specific product – for example N95 masks.

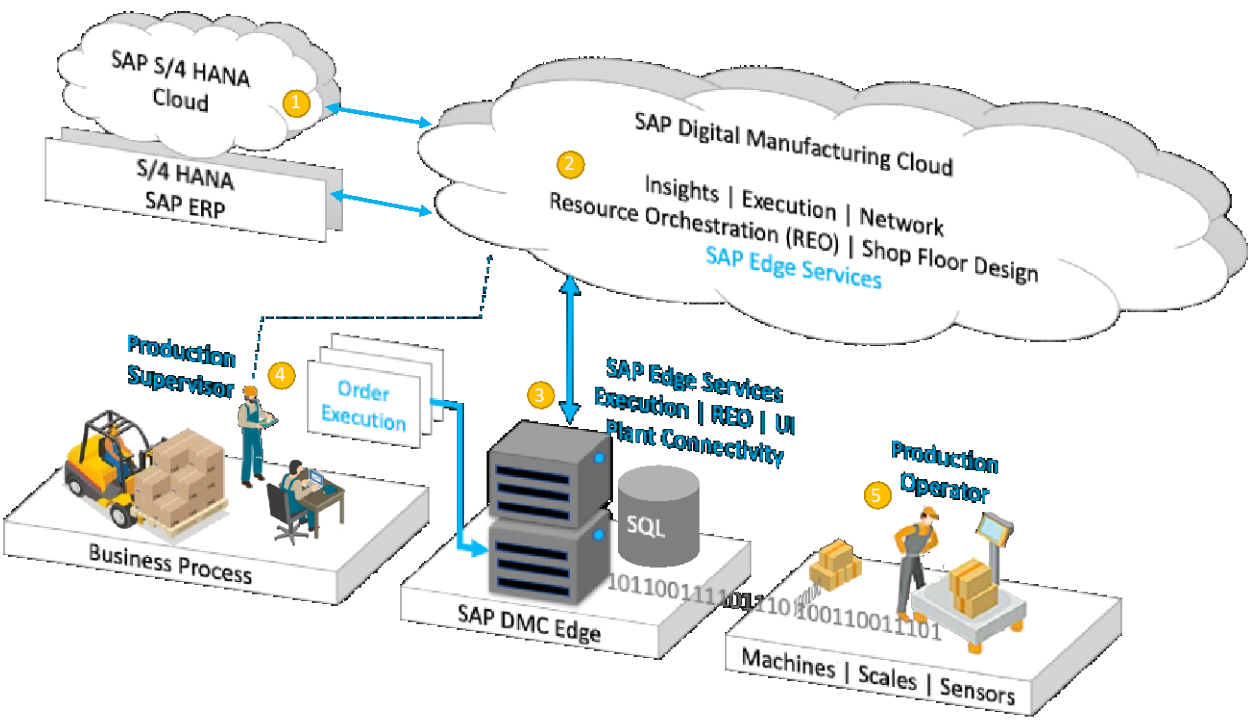

Fig. 7 - Scenario of mission critical manufacturing process using SAP's edge computing solutions.

In the first step of Figure 7, the incoming order that is initially placed in S/4 HANA Cloud is integrated via standards interfaces into Digital Manufacturing Cloud (DMC). DMC covers functionalities for production process design, resource orchestration, manufacturing execution, monitoring, and analytics. Using its business functionalities, DMC ensures through configurable data partitioning that this order and its associated data are routed only to the correct plant. Such central cloud plane ensures low TCO and oversight across a potentially complex edge topology. The DMC for edge computing deployed at this plant obtains the order for N95 masks from DMC in the cloud. Both DMC in the cloud and at the edge are seamlessly integrated considering security standards.

Production Supervisor on the plant is responsible for managing the production lines, the staff assigned to this plant, and materials. Last minute factors such as workers calling sick, machine breakdowns, and unavailability of materials and resources, require prompt and agile changes to the production line. Consider that the next order in the production line is the N95 masks, but the quality measurements tools needed to fulfill the quality standards of N95 are unavailable due to faults. The *Production Supervisor*, using DMC for edge computing, postpones this Order Execution and starts a run of consumer grade masks instead, even though the internet is not available. Business functionalities available in DMC for edge computing such as Business Data Partitioning and Business Object Synchronization allow that the respective business objects such as production order, execution order, and resource orchestration have been already synchronized as cloud connectivity was available. This ensures delays are avoided and keep to their production targets (e.g., keep OEE as high as possible).

Also, on the plant, *Production Operators* use DMC for edge computing to run the order and create confirmations while completing the order. As shown in Figure 7, data-driven insights are enabled with IoT components. Plant connectivity connects to machines and ingest data using business context to identify and illustrate quality or non-conformance with a production order. Confirmations are then updated and associated with the order at the edge in real time.

By enabling edge deployment of cloud capabilities, customers can add resilience to their mission critical business processes when unforeseen last-minute changes can be handled locally and quickly.

DISCUSSION & OUTLOOK

Discussion

This paper discussed how cloud orchestration solutions can drive Industry 4.0 at distributed edges and introduced the key requirements of modern edge computing approaches to enable mission critical execution. We outlined recommendations on what business applications and data are needed as well as where to deploy them within the edge-cloud continuum to ensure the customer specific business value proposition. Edge computing solutions such as SAP Digital Manufacturing Cloud for edge computing and Edge Services provide a full set of features enabling mission critical execution of distributed manufacturing processes. These enterprise-ready solutions enable the edge deployment of cloud capabilities. By bringing capabilities from cloud native to the edge and orchestrating it from the cloud, enterprise solutions can add the following value to business processes:

- **Flexibility:** Applications can be deployed, orchestrated, and managed on edge devices with a variety of processing and storage capabilities (light (small), medium, heavy).

- **Scalability:** Deployment and configuration of business applications are individualized according to customer-specific need and scenario as well as connectivity availability.
- **High Availability:** Capabilities of Kubernetes can be leveraged to create High Availability configurations for edge computing deployments. High availability configurations optimize response times and minimize downtime and are crucial in mission critical scenarios.
- **Faster response times:** Edge computing enables real-time decision-making capabilities within manufacturing processes to ensure business continuity and even if intermittent connectivity, limited bandwidth, and latency occurs.
- **Robustness:** Increased resilience to mission critical business processes since unforeseen last-minute changes can be handled locally and quickly.
- **Low TCO:** Lower TCO via simplified lifecycle management and orchestration of high complex topologies edge nodes from the cloud.
- **Better customer experience**: Optimal balance between standardized components and increased customer experience via simplified configuration of business application for specific scenarios.
- **Plug and play:** The consideration of industry standards within solutions for edge computing bring several benefits - *plug and play* via out of the box interoperability, increased acceptance of new technologies, and lower integration and operational costs.

Outlook

Although edge computing solutions provide key building blocks to enable mission critical execution of distributed processes across the supply chain, the edge computing technology still maturing, and several concepts can be explored in future works. The following points provide a preview of future developments in context of edge computing and Industry 4.0:

- **De-facto standards for highly distributed and heterogeneous edge computing landscapes:** The establishment of de-facto standards to address the heterogeneous edge computing landscape will be decisive to avoid lock-in effects and node to node interoperability.
- **Asset Administration Shell and edge computing:** The question on how to leverage AAS in edge computing infrastructures to support digital twin across all phases of Design to Operate (D2O) will be addressed in future works. Also, the question on how edge computing and AAS could be explored to combine data from multiple sources and different systems (IT and OT) and to ensure the continuity of digital twins in highly distributed scenarios must be discussed. Since AAS would bring intelligence to assets, one argumentation to be evaluated is the convergence of the IoT assets layer into the Edge layer from the edge-cloud continuum. Using the AAS, some business applications and data could be already consumed at the asset. Such approach must be discussed in future works since this will have an impact on the shop floor and other participants of the supply chain.

  Edge Computing as enabler of data sovereignty: Deeper discussions on how edge computing addresses data sovereignty requirements in context of Gaia-X and Catena-X are needed. As the activities within Gaia-X and Catena-X are just beginning, it is necessary to assess on how their outcomes could affect current edge computing solutions. Another question to be evaluated is how to leverage existing concepts by bringing business applications and data closer to the data source in a data space consisting of distributed plants from different companies.

References

[1] E. Arora, "Design to Operate: Run a Resilient, Sustainable Supply Chain," SAP SE, 12 2020. [Online]. Available: https://blogs.sap.com/2020/12/01/design-to-operate-run-a-resilient-sustainable-supply-chain/. [Accessed 05 2021].

[2] SAP SE, "SAP Enabling Technologies for Industry 4.Now," 09 2020. [Online]. Available: https://www.sap.com/documents/2020/09/704c5ac8-b17d-0010-87a3-c30de2ffd8ff.html. [Accessed 05 2021].

[3] SAP SE, "Why the Time for Industry 4.0 Is Now," 06 2020. [Online]. Available: https://news.sap.com/2020/06/time-industry-4-now/. [Accessed 05 2021].

[4] IDC, "IDC FutureScape: Worldwide IoT 2020 Predictions," IDC, 2019.

[5] Linux Foundation, "LF Edge - Data At The Edge," 09 2019. [Online]. Available: https://www.stateoftheedge.com/reports/data-at-the-edge-2019/.

[6] A. Bhatia, D. Litovsky, A. Drewery, N. Misra and Z. Yusuf, "The Battle at Computing's Edge," 03 2021. [Online]. Available: https://www.bcg.com/de-de/publications/2021/what-it-takes-to-win-the-edge-computing-battle. [Accessed 05 2021].

[7] Linux Foundation, "Edge Networking: An Introduction," 2021. [Online]. Available: https://www.lfedge.org/wp-content/uploads/2020/06/EdgeNetworkingIntroduction_Whitepaper.pdf.

[8] M. Tseng, T. Edmunds and L. Canaran, "Introduction to Edge Computing in IIoT: An Industrial Internet Consortium White Paper," 06 2018. [Online]. Available: https://www.iiconsortium.org/pdf/Introduction_to_Edge_Computing_in_IIoT_2018-06-18-updated.pdf. [Accessed 05 2021].

[9] J. Zao, C. Byers, B. Murphy, S. AbiEzzi, D. Banks, K. An, F. Michaud and K. Bartfai-Walcott, "The Industrial Internet of Things Distributed Computing in the Edge," 10 2020. [Online]. Available:

https://www.iiconsortium.org/pdf/IIoT-Distributed-Computing-in-the-Edge.pdf. [Accessed 05 2021].

[10] A. Willner and V. Gowtham, "Toward a Reference Architecture Model for Industrial Edge Computing," *IEEE Communications Standards Magazine,* vol. 4, no. 4, pp. 42-48, 2020.

[11] K. Cao, S. Hu, Y. Shi, A. Colombo, K. Cao and S. Karnouskos, "A Survey on Edge and Edge-Cloud Computing Assisted Cyber-Physical Systems," *IEEE Transactions on Industrial Informatics.*

[12] Linux Foundation, "Sharpening the Edge: Overview of the LF Edge Taxonomy and Framework," 07 2020. [Online]. Available: https://www.lfedge.org/wp-content/uploads/2020/07/LFedge_Whitepaper.pdf. [Accessed 08 2020].

[13] Eclipse Foundation, "Open Source Software for Industry 4.0," 2017. [Online]. Available: https://iot.eclipse.org/community/resources/white-papers/pdf/Eclipse%20IoT%20White%20Paper%20-%20Open%20Source%20Software%20for%20Industry%204.0.pdf. [Accessed 05 2021].

[14] SAP SE, "Setup and Operations Guide for SAP Digital Manufacturing Cloud for edge computing - Basic Concepts," 2021. [Online]. Available: https://help.sap.com/viewer/Edge%20Setup%20and%20Operations%20Guide. [Accessed 05 2021].

[15] SAP SE, "GitHub Repository SAP i40-aas," [Online]. Available: https://github.com/SAP/i40-aas. [Accessed 05 2021].

[16] SAP SE, "SAP Digital Manufacturing Cloud," [Online]. Available: https://www.sap.com/products/digital-manufacturing-cloud.html?infl=2a02ae07-5d8c-4777-a55a-9c83fcdf1c9f. [Accessed 05 2021].

[17] SAP SE, "SAP Business Technology Platform: The Platform for the Intelligent Enterprise," 2021. [Online]. Available: https://www.sap.com/products/digital-manufacturing-cloud.html?infl=2a02ae07-5d8c-4777-a55a-9c83fcdf1c9f. [Accessed 05 2021].

[18] SAP SE, "Setup and Operations Guide for SAP Digital Manufacturing Cloud for edge computing: Operations Overview," [Online]. Available: https://help.sap.com/viewer/Edge%20Setup%20and%20Operations%20Guide/02bac5b5631d4d2da4f0d4c71e045a1f.html. [Accessed 05 2021].

[19] SAP SE, "Edge Services Overview Guide," [Online]. Available: https://help.sap.com/viewer/17e9e088c03b4d0cb59cd848e4e1c886/2105/en-US/493b8ad6c8c04f38b32c380524e01c20.html. [Accessed 05 2021].

[20] SAP SE, "Setup and Operations Guide for SAP Digital Manufacturing Cloud for edge computing: Architecture and Component Overview," [Online]. Available: https://help.sap.com/viewer/Edge%20Setup%20and%20Operations%20Guide/c78bde5411da497b9e6d2509a045d49b.html.

[21] SAP SE, "Edge Services Overview Guide: Architecture for SAP IoT Integration," [Online]. Available: https://help.sap.com/viewer/17e9e088c03b4d0cb59cd848e4e1c886/2105/en-US/eb607054edf548b29a309f5955f50afe.html. [Accessed 05 2021].

[22] Gardener, "Gardener - Deliver fully-managed clusters at scale everywhere with your own Kubernetes-as-a-Service," [Online]. Available: https://gardener.cloud.

[23] R. Franzke and V. Chandrasekhara, "Gardener - The Kubernetes Botanist," Kubernetes, 05 2018. [Online]. Available: https://kubernetes.io/blog/2018/05/17/gardener/. [Accessed 05 2021].

[24] Standardization Council Industrie 4.0, "German Standardization Roadmap Industrie 4.0," 03 2020. [Online]. Available: https://www.sci40.com/sci-4-0/normungsroadmap/.

[25] E. Barnstedt, B. Boss, E. Clauer, D. Isaacs, S.-W. Lin, M. Somayeh, P. van Schalkwyk and T. Weber Martins, "Open Source Drives Digital Twin Adoption," *IIC Journal of Innovation: Innovations in Digital Twins,* no. 16, pp. 19-34, 03 2021.

[26] OPC UA Foundation, "OPC UA Foundation," [Online]. Available: https://opcfoundation.org. [Accessed 06 2021].

[27] https://github.com/OPCFoundation, "GitHub repository of OPCFoundation," [Online]. Available: https://github.com/OPCFoundation. [Accessed 06 2021].

[28] Eclipse Foundation, "Eclipse IoT Working Group," [Online]. Available: https://iot.eclipse.org. [Accessed 06 2021].

[29] Eclipse Foundation, "Edge Native WG," [Online]. Available: https://edgenative.eclipse.org.

[30] Eclipse Foundation, "From DevOps to EdgeOps: A Vision for Edge Computing," 2021. [Online]. Available: https://outreach.eclipse.foundation/edge-computing-edgeops-white-paper. [Accessed 06 2021].

[31] Linux Foundation, "LF Edge," [Online]. Available: https://www.lfedge.org. [Accessed 06 2021].

[32] OpenInfra Foundation, "OpenInfra Foundation," [Online]. Available: https://openinfra.dev. [Accessed 06 2021].

[33] B. Balter, "Why Open Source," [Online]. Available: https://ben.balter.com/2015/11/23/why-open-source/.

[34] Kyma, "Kyma - Extend your applications with Kubernetes," [Online]. Available: https://kyma-project.io.

[35] J. Cawley, "What is Kyma?," SAP SE, 07 2019. [Online]. Available: https://blogs.sap.com/2019/07/08/what-is-kyma/.

[36] B. Boss, S. Malakuti , S.-W. Lin , T. Usländer , E. Clauer, T. Usländer , M. Hoffmeister and L. Stojanovic , "Digital Twin and Asset Administration Shell Concepts and Application in the Industrial Internet and Industrie 4.0 - An Industrial Internet Consortium and Plattform Industrie 4.0 Joint Whitepaper," 09 2020. [Online]. Available: https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/Digital-Twin-and-Asset-Administration-Shell-Concepts.pdf?__blob=publicationFile&v=11. [Accessed 05 2021].

[37] Plattform Industrie 4.0, "Details of the Asset Administration Shell, Part 1 - The exchange of information between partners in the value chain of Industrie 4.0," 11 2020. [Online]. Available: https://www.plattform-i40.de/PI40/Redaktion/EN/Standardartikel/specification-administrationshell.html. [Accessed 05 2021].

[38] Plattform Industrie 4.0, "Details of the Asset Administration Shell, Part 2," 11 2020. [Online]. Available: https://www.plattform-i40.de/PI40/Redaktion/EN/Standardartikel/specification-administrationshell.html. [Accessed 05 2021].

[39] Industrial Digital Twin Association, [Online]. Available: https://idtwin.org/en/. [Accessed 05 2021].

[40] Industrial Digital Twin Association, "The Digital Twin for Industry 4.0," 03 2021. [Online]. Available: https://idtwin.org/wp-content/uploads/2021/03/IDTA_AAS_de_03-01_pub_en.pdf. [Accessed 05 2021].

[41] ZVEI (German Electrical and Electronic Manufacturers' Association), "ExamplesoftheAssetAdministrationShellforIndus- trie 4.0 Components," 04 2017. [Online]. Available: https://www.zvei.org/fileadmin/user_upload/Presse_und_Medien/Publikationen/2017/April/Asset_Administration_Shell/ZVEI_WP_Verwaltungschale_Englisch_Download_03.04.17.pdf. [Accessed 05 2021].

[42] IDSA, "International Data Spaces Association - The future of the data economy is here," [Online]. Available: https://internationaldataspaces.org. [Accessed 05 2021].

[43] International Data Spaces Association, "IDSA Reference ARchitecture Model," 04 2019. [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf. [Accessed 05 2021].

[44] Plattform Industrie 4.0, "Creating the DataSpace Industrie 4.0," 04 2021. [Online]. Available: https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/PositionPaper-DataSpace.pdf?__blob=publicationFile&v=5. [Accessed 05 2021].

[45] International Data Spaces Association, "GitHub repository International Data Spaces Association," [Online]. Available: https://github.com/International-Data-Spaces-Association.

[46] Open Manufacturing Platform, "Insights Into Connecting Industrial IoT Assets," 12 2020. [Online]. Available: https://open-manufacturing.org/wp-content/uploads/sites/101/2020/12/OMP-IIoT-Connectivity-White-Paper-20201207.pdf.

[47] Open Manufacturing Platform, "GitHub Repository Insights into Connecting Industrial IoT Assets White Paper," [Online]. Available: https://github.com/OpenManufacturingPlatform/iot_connectivity_public/blob/publication/White_Paper/01_Insights_Into_Connecting_Industrial_IoT_Assets/00_Acknowledgements.md.

[48] Gaia-X, "Gaia-X - A Federated and Secura Data Infrastructure," [Online]. Available: https://gaia-x.eu. [Accessed 05 2021].

[49] SAP SE, "Catena-x - The Alliance for Secure and Cross-Company Data Exchange in the Automotive Industry Is Picking up Speed," 02 2021. [Online]. Available: https://news.sap.com/2021/03/catena-x-automotive-alliance-picks-up-speed/. [Accessed 05 2021].

[50] The Wall Street Journal, "Companies Extend Cloud to the Edge," 03 05 2020. [Online]. Available: https://www.wsj.com/articles/companies-extend-cloud-to-the-edge-11620080377. [Accessed 05 2021].

[51] M. Grieves, "Digital twin: Manufacturing excellence through virtual factory replication," 2014. [Online]. Available: http://www.apriso.com/library/Whitepaper_Dr_Grieves_DigitalTwin_ManufacturingExcellence.php.

ACKNOWLEDGEMENTS

The views expressed in the *IIC Journal of Innovation* are the contributing authors' views and do not necessarily represent the views of their respective employers nor those of the Industrial Internet Consortium.

➢ Return to IIC Journal of Innovation landing page for more articles and past editions