The process data twin for digitized production

Andreas Leinenbach, Marcel Wagner, Fabio Oettl, Johannes Schilp

Angaben zur Veröffentlichung / Publication details:

Leinenbach, Andreas, Marcel Wagner, Fabio Oettl, and Johannes Schilp. 2024. "The process data twin for digitized production." In *Proceedings of the CPSL 2024 - 6th Conference on Production Systems and Logistics*, edited by D. Herberger and M. Hübner, 647–55. Hannover: publish-Ing. https://doi.org/10.15488/17753.

6th Conference on Production Systems and Logistics

The Process Data Twin For Digitized Production

Andreas Leinenbach[1], Marcel Wagner[2] , Fabio Oettl[3], Johannes Schilp[1]

[1]Lehrstuhl für Produktionsinformatik, Universität Augsburg, Germany
[2]Intel Deutschland GmbH, Feldkirchen, Germany
[3]IndustryFusion Foundation, St. Gallen, Switzerland

Abstract

In this paper, we describe the Process Data Twin (PDT) of the Industry Fusion Foundation (IFF). IFF has been founded by small and medium manufacturers in Europe to build open-source solutions for digitized production. The PDT is a data-driven manufacturing platform that can link and contextualize data from multiple sources. In addition, it is using World Wide Web Consortium semantic standards for vendor-independent data exchange between machines, software, and other assets in a factory. The Core of the PDT is a context broker (CB), as standardized by the European Telecommunications Standards Institute. The CB handles the communication between different assets and models and controls production processes. Every asset is registered and assigned a globally unique identifier by an IFF-operated service. This allows optimal collaboration and data sharing between the members of IFF. We will show how the whole solution simplifies and optimizes production processes, especially when aspects of processes need to be shared between parties.

Keywords

Digital Twin; Process Data Twin; Conference; CPSL

1. Introduction

The term "digital twin" has become increasingly important in recent years, particularly in the context of Industry 4.0 and the Internet of Things (IoT). In the manufacturing industry, the digital twin is often used to optimize production processes, minimize downtime and improve quality. Simulation allows companies to identify and eliminate potential problems before they occur in real-life operation. In the area of plant management, the digital twin enables precise monitoring and maintenance of plants and infrastructures. By continuously analysing operating data, deviations can be detected at an early stage and proactive measures can be taken to prevent failures and increase efficiency.

The Process Data Twin attempts to think about production from the perspective of the production process and to collect the required data. The focus is primarily on process and production parameters and less on the company's machine view. The Process Data Twin, the selection of components used and the associated communication infrastructure form a functioning overall system and thus differ from selective communication solutions such as OPC-UA or passive data collections such as the asset administration shell.

The paper first describes the current state of science and in particular the asset administration shell and, with OPA-UA, an exemplary communication interface, as well as the technological basis for the implementation of the Process Data Twin. This is followed by the introduction of the process data twin from a technical and organizational perspective, with the constraints within the PDT being a key feature.

DOI: https://doi.org/10.15488/17753
ISSN: 2701-6277

2. State of Research

2.1 Digital Twin

There are various definitions of the term "digital twin". In [1], the digital twin is described as "a digital representation of a physical object or system that is continuously synchronized with the real object and thus forms a bidirectional information bridge between the physical and virtual worlds.". Further definitions, according to [2], speak of a virtual representation of an object or process that [...] is updated in real time. Kritzinger describes different levels of abstraction of digital twins and considers communication to be the central distinguishing feature between the real object and the digital representation [3]. If there is no automatic communication between the physical instance and the digital representation, this is referred to as a digital model. If there is automatic communication in the digital image, this is referred to as a digital shadow. In the case of automatic, bidirectional communication from the physical instance towards the digital representation, it is defined as a digital twin.

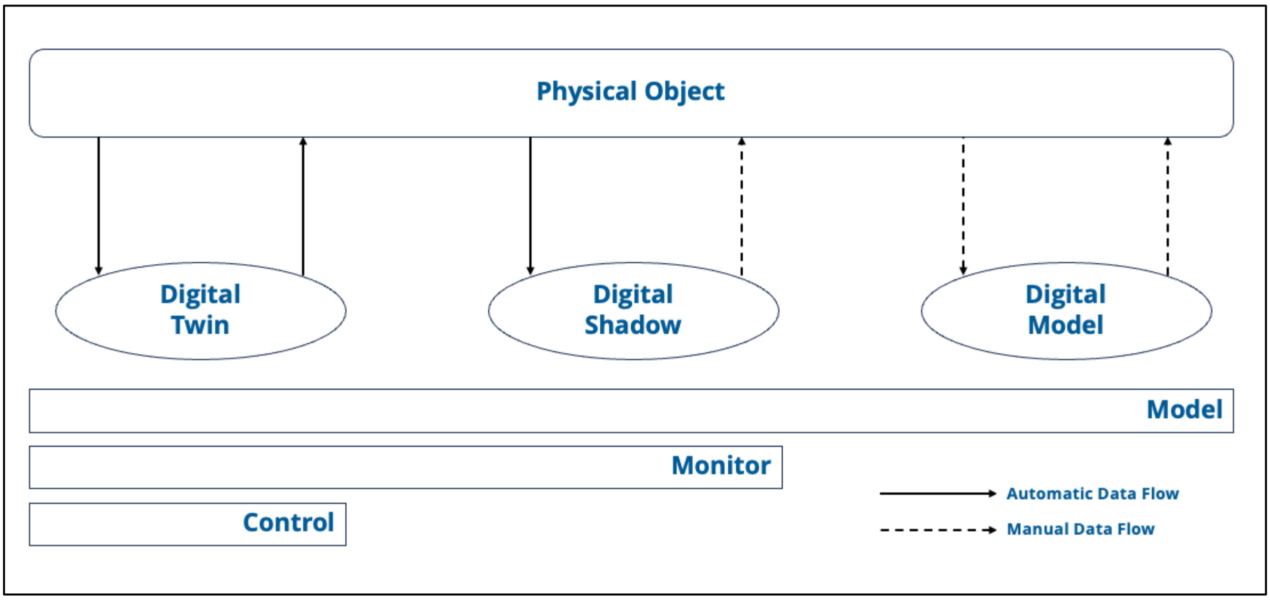

Figure 1: Interaction between the physical object and the digital representative [3]

2.2 Asset Administration Shell

The asset administration shell (AAS) is a concept in the field of Industry 4.0 and acts as a digital representation of a physical object or entity. It contains all relevant information about the object, which can be both static and dynamic data from the process. The information storage is standardized and thus enables interoperability between different systems via a passive API. In [4], the AAS is described as a central interface for accessing and managing information. [5] It can serve as a database for integrating data from different sources and can therefore be used for analyses. In addition, access rights and protection mechanisms can be integrated to prevent unauthorized data access. However, active integration of the AAS via the passive API provided is necessary in order to be able to use it in an industrial environment. [6]

2.3 OPC UA

The Open Platform Communications Unified Architecture (OPC UA) is a communication technology that is now used by many machine manufacturers for networking and connecting machines. [7] OPA UA offers the possibility of standardizing uniform information models in various data structures. In addition, different communication patterns are supported, such as point-to-point, publish/subscribe or client/server relationships with various protocols such as TCP, UDP or higher-level abstractions such as MQTT. Various protection mechanisms, such as encryption, authentication and authorization, secure the communication against unauthorized access. OPC UA is standardized in IEC 62541. [8]

2.4 JSON-LD & NGSI-LD

JavaScript Object Notation for Lined Data (JSON-LD) provides structured data in a standardized format. JSON-LD is based on the linked data principle, which classifies data and thus makes it readable in a standardized way. [9] By using special keywords and conventions, entities and their relationships can be clearly identified and described in a JSON document. This enables different applications to interpret the data and create semantic links between different resources. Next Generation Service Interface for Linked Data (NGSI-LD) is a standard of the European Telecommunications Standards Institute (ETSI). NGSI-LD builds on the previous NGSI standard, which was mainly used for the management of contextual information in smart city environments. NGSI-LD extends this approach to facilitate the interoperability and integration of data across different domains. It is based on the Linked Data principle and uses the JSON-LD format for the representation of data. While JSON-LD is a general method for semantic annotation of JSON data, NGSI-LD is more specific to the Internet of Things (IoT) domain. [10]

```
{
  "@context": {
    "name": "http://schema.org/name",
    "birthDate": "http://schema.org/birthDate",
    "address": "http://schema.org/address"
  },
  "@type": "Person",
  "name": "Max Mustermann",
  "birthDate": "2000-01-01",
  "address": {
    "@type": "PostalAddress",
    "addressLocality": "Musterstadt",
    "addressCountry": "Germany"
  }
}
```

```
{
  "id": "urn:ngsi-ld:Person:1234",
  "type": "Person",
  "name": {
    "type": "Property",
    "value": "Max Mustermann"
  },
  "birthDate": {
    "type": "Property",
    "value": {
      "@type": "DateTime",
      "@value": "2000-01-01T10:00:00Z"
    }
  },
  "address": {
    "type": "Relationship",
    "object": {
      "id": "urn:ngsi-ld:PostalAddress:5678",
      "type": "PostalAddress",
      "addressLocality": {
        "type": "Property",
        "value": "Musterstadt"
      },
      "addressCountry": {
        "type": "Property",
        "value": "Germany"
      }
    }
  }
}
```

Figure 2: Comparison between JSON and JSON-LD

2.5 SHACL

Shapes Constraint Language (SHACL) is a standardized language for describing data schemas and constraints in the Semantic Web. Developed by the World Wide Web Consortium (W3C), SHACL enables the definition and verification of structured data models to improve the quality and interoperability of data on the web. The SHACL specification defines a variety of constructs for describing data forms, properties and constraints. By using SHACL, precise rules and restrictions can be defined that must be adhered to by data instances. These rules are defined as so-called "shapes", which specify the expected properties and structures of the data. An important feature of SHACL is its flexibility and expressiveness. It supports a variety of data structures and types, including simple and complex data types. In addition, it provides mechanisms for linking shapes, defining hierarchies and reusing rules, which facilitates the modeling of complex data structures.[11] [12]

3. Process Data Twin

The Process Data Twin (PDT) is embedded in an overall system that is currently being developed as an open-source solution by the IndustryFusion Foundation. Therefore, production is viewed from a process perspective rather than the individual machine and its digital representation as a model. Both the technical approach and the structural design are described below and the communication structure is discussed.

3.1 Technical Approach

The PDT is an analytics system designed for low-latency operations. It integrates several technologies to manage and analyze data effectively.The ScorpioBroker, a key component of the PDT, functions as the object or context broker. It employs NGSI-LD, an extension of JSON-LD, to manage relationships between entities and append metadata, such as timestamps, to these entities. NGSI-LD serves dual purposes as a data specification and an API for interacting with entities. Data is streamed to the platform via MQTT, a messaging protocol for small sensors and mobile devices, or through the NGSI-LD Representational State Transfer Application Programming Interface (REST-API), a software architectural style that sets constraints for creating Web services. Kafka, a distributed streaming platform, acts as the central message bus within the system. It is used for building real-time data pipelines and streaming applications. For stateful analytics, the system uses Flink, a framework and distributed processing engine. The system translates SHACL rules and the knowledge graph into StreamingSQL, which are then executed in real-time. This process ensures data compliance with specific rules or patterns and enhances the semantic understanding of the data.

The open-source implementation of the PDT is available at Github. [13] This system provides a structured and semantically rich approach to data management, representation, and analysis.

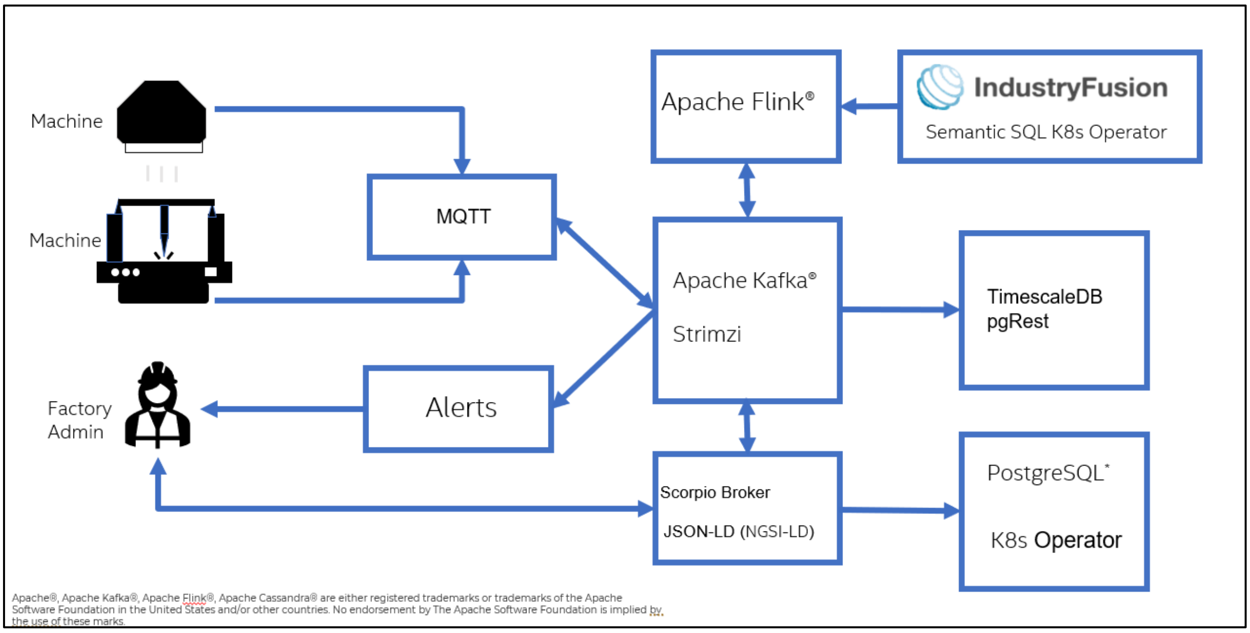

Figure 3: Technical Structure of PDT

3.2 Constraints in the PDT

In the PDT, data objects and the knowledge graph can contain arbitrary data and links. To manage the integrity of data, we use SHACL to define data constraints. One of the key advantages of using SHACL is its independence from its implementation. This means that constraints can be defined and checked without being tied to a specific implementation, and these constraints can even be shared across different implementations. Furthermore, the declarative nature of SHACL simplifies the process of maintaining an overview and scaling the system compared to a procedural implementation. This is because SHACL allows us to express what we want to achieve (the "what") rather than detailing the steps to achieve it (the "how").

To illustrate, consider the following examples. The exemplary assets labeled with ‘ex:’ are being considered in this regard. In example one, shown in figure 4, the combination of a plasma cutter and a filter machine is considered. Technologically, it is imperative, that each plasma cutter has a corresponding filter machine to extract process gases. We can use SHACL to check if an object of type 'ex:plasmacutter' has a relationship 'ex:hasFilter' and is linked by that to an existing object of type 'ex:filter'. This realtion is manufacturer-independent and thus applicable at any time.

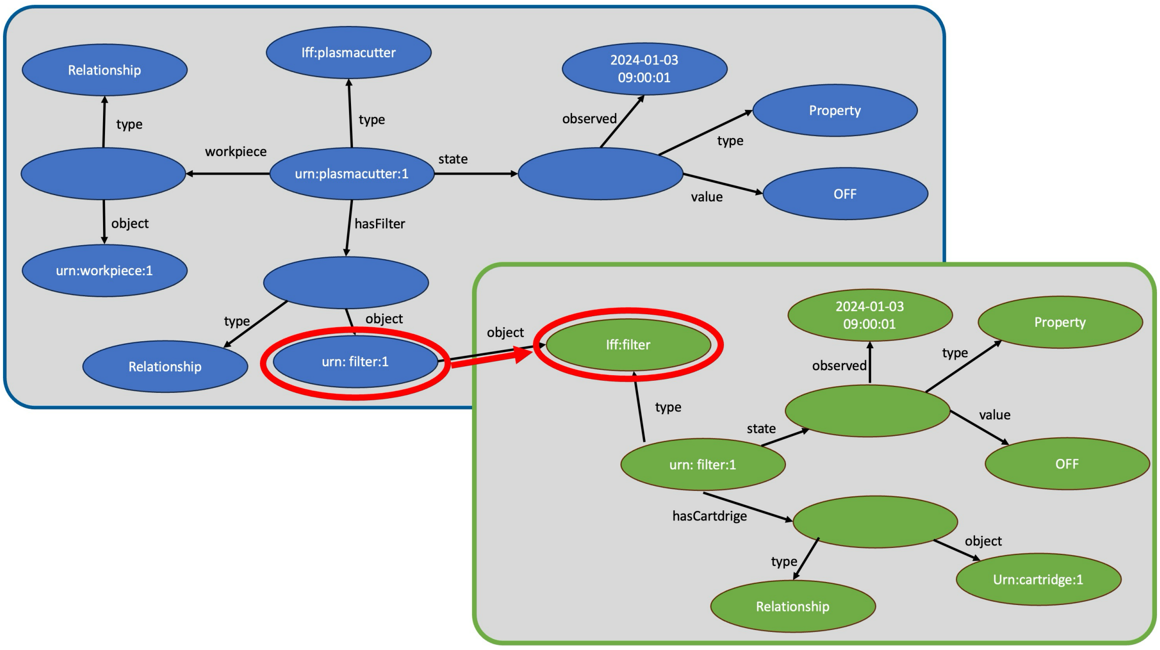

Figure 4: Constraints in the PDT

In the second example, the dynamic operation of machinery between the filter system and plasma cutter is considered. We can also use SHACL to check if the object 'ex:plasmacutter' is running with state 'ex:state_PROCESSING', the linked filter must be in state 'ex:state_ON'. In summary, the use of SHACL in the PDT provides a flexible and scalable approach to constraint management, allowing us to define and check constraints independently of the implementation.

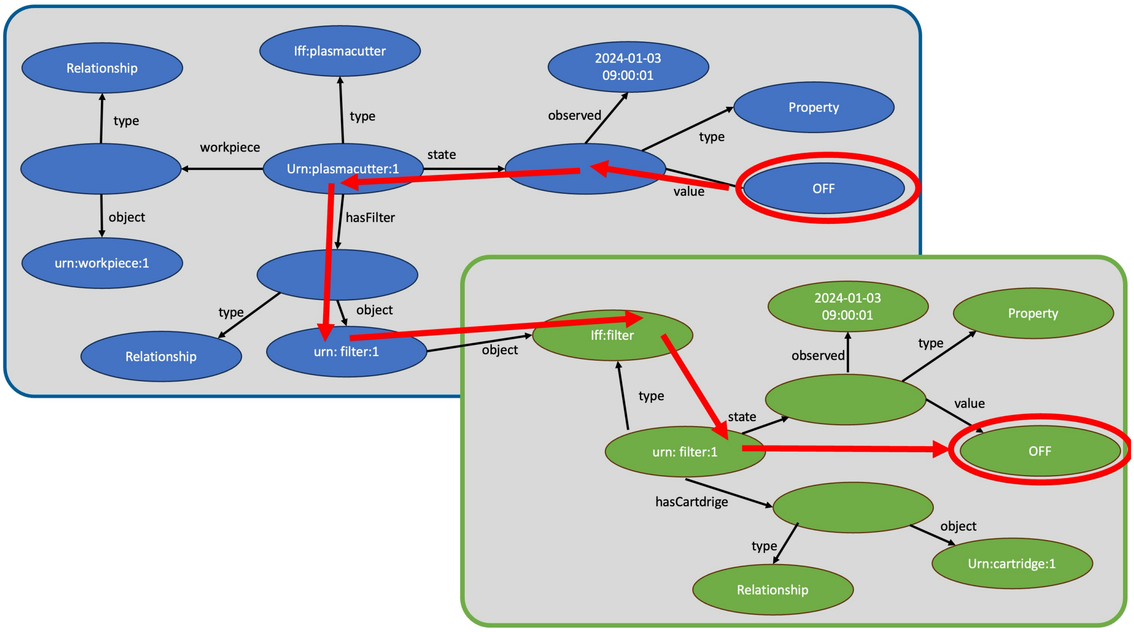

Figure 5: Second Example for Constraints

3.3 Structural Approach

Every asset has a PDT. Assets can be machines, logistics vehicles, software components or people. The PDT holds all the necessary static and dynamic information and optionally, it contains the globally unique identification, which enables information, production data and documents to be assigned. The existing data can be structured in four areas shown in figure 6, whereby the technical implementation is carried out using the methods described, such as JSON-LD and the context broker.

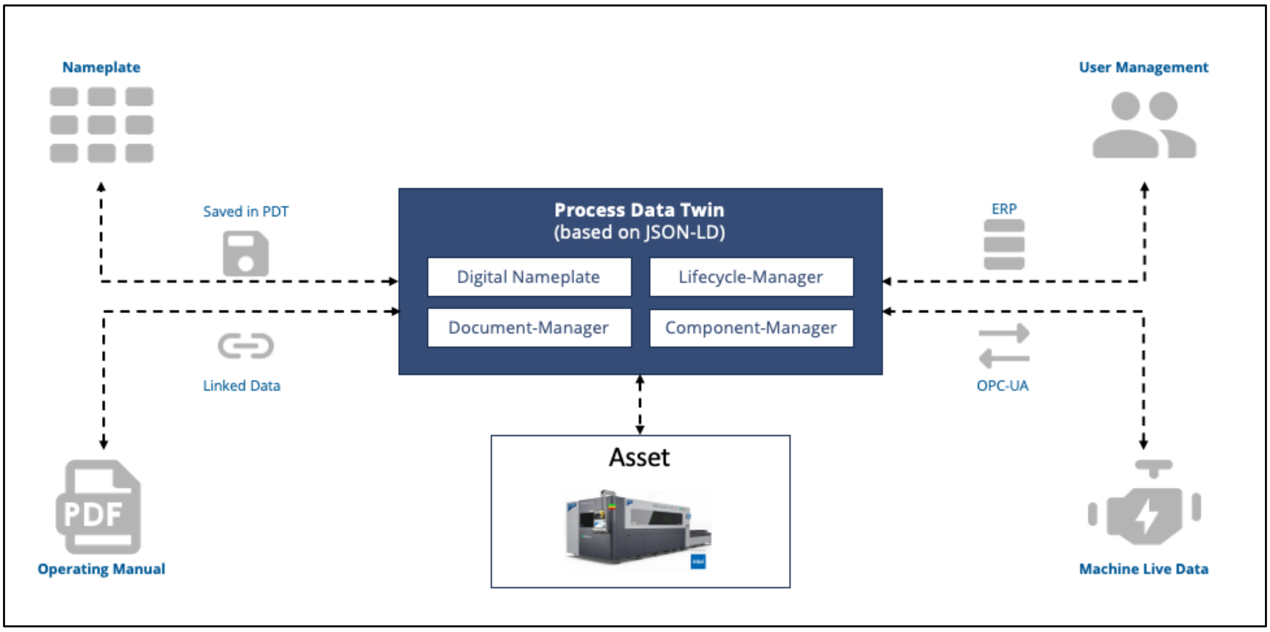

Figure 6: Structural view on PDT

The digital name plate contains the classic, basic machine or asset information. In addition to the manufacturer, this includes the machine type, the year of manufacture and the owner or installation information. The Document Manager provides documents such as CE certification, operating documents or maintenance plans. An active interface to other assets, such as an ERP system, can be used to retrieve proof and authorizations for operating machines and check the authorization of employees on the respective assets. Current and past operating data can be provided, stored and evaluated via typical communication interfaces such as OPC-UA or MQTT. The semantic structure of JSON-LD allows manufacturer-independent evaluation of data and the development of universal AI models and algorithms.

The entire ecosystem can be seen in Figure 7. It consists of three levels. The lowest level contains the assets that interact within the system. It does not matter whether this is a software system, a machine or a storage system. These assets communicate in a standardized way via MQTT to the middle level, where the factory server is located. Information can be stored there, but it can also be evaluated directly. Manufacturer-independent communication and data storage make it possible to create and check models and relationships between assets without specifying a manufacturer-specific definition. Basic data, such as the machine status in accordance with VDI 3423, is provided by each asset and thus offers the option of using the Factory Manager to keep track of the current production status. The Factory Manager is an application that runs at the edge level and displays the basic information of the connected machines, e.g. energy consumption or machine status. In this environment, it is possible to connect additional machines to the Edge Server so that they are immediately available on the Factory Server and, if necessary, for other services without further integration measures. This is also known as Plug and Produce.

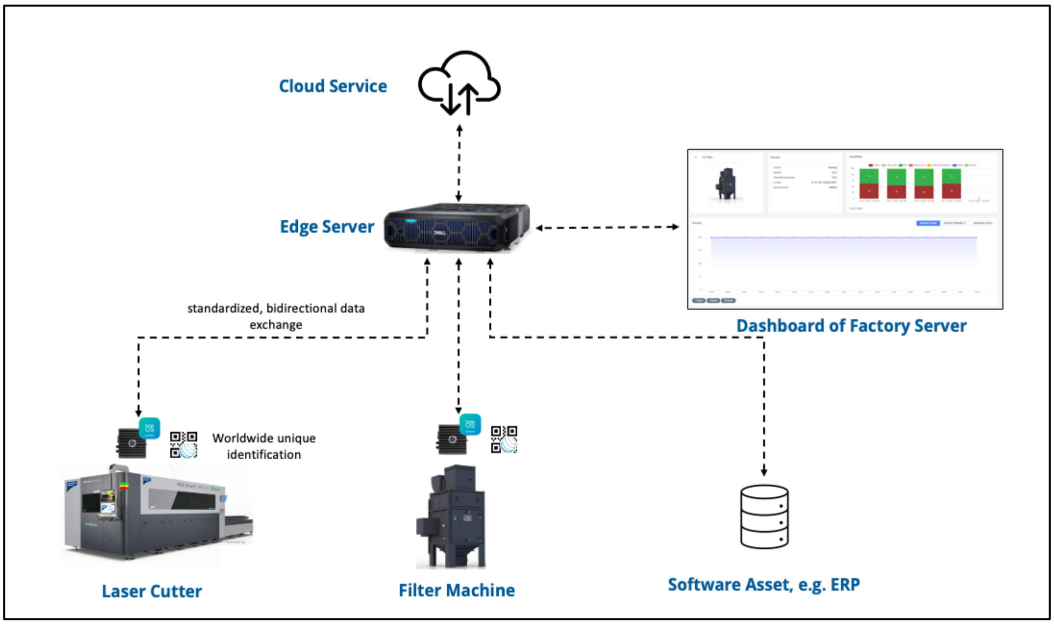

Figure 7: Ecosystem for PDT

Cloud services are used for cross-company communication, enabling the networking of machine manufacturers and users, as well as the implementation of production networks and applications such as cloud manufacturing. Two use cases are described as examples in chapter 4.

In the context of AAS and OPC-UA, the PDT offers a universal, holistic solution for realizing digital twins without additional implementation effort, such as addressing the passive API of the AAS or evaluating the information models of the OPC-UA interface. At the same time, the AAS and OPC UA can act as part of the ecosystem and use and integrate the existing, standardized infrastructure.

4. Use Cases

In this section, two use cases that are made possible by the application of PDT are presented as examples. In addition, the benefits for companies resulting from the implementation of the use cases are presented. These were implemented by companies that are part of the IndustryFusion Foundation. The IFF is an association of several medium-sized companies with the aim of achieving end-to-end, manufacturer-independent networking of machines. The open source developments are regularly presented at various events. [15]

4.1 Lasercutter & Air Filter

Saving energy is an important factor in the manufacturing industry. One use case that could contribute to energy savings is the exchange of data between laser cutters and air filters. Cutting a metal plate produces metal vapors that pose a health risk to people and the environment. Therefore, the task of the air filter is to generate sufficient negative pressure so that the resulting vapors are sucked in. The required negative pressure is heavily dependent on the dimensions of the metal plate, as it covers more of the laser cutter's processing table depending on the plate's size. Very large metal plates cover the guide table almost completely, which means that less air needs to be sucked in to generate the required vacuum. Small metal plates, on the other hand, only partially cover the processing table, which means that more air has to be drawn in.

It is currently common practice for the air filter to always run at full power to ensure that sufficient air is drawn in, as it has no knowledge of the components to be cut. It is technically possible to regulate the fan power. However, the control must ensure that there is sufficient negative pressure even with reduced fan power so that the vapors produced do not enter the factory and thus do not endanger the safety of people and the environment.

The use of a PDT enables the standardized exchange of semantic data between different assets in a factory. This allows the relevant information about the dimensions of the metal plate to be cut and to be exchanged. In addition to the length and width of the metal plate, the thickness is also crucial. With a thicker metal plate, the laser removes more material and therefore produces more vapors. This information is exchanged between the laser cutter and the air filter via the PDTs of the two assets, significantly reducing the filter's energy consumption.

4.2 Continous Production Process

The use of PDT in the embedded overall structure enables the end-to-end monitoring and tracking of a production process. By identifying and collecting data from both the component and the assets involved, such as ERP, machines and employees, the entire process can be traced and documented. Regardless of whether the system actually uses the PDT, it stores important information on the source material, the course of the material and the effects of production processes such as mechanical or thermal influences. In the further course of the process, both the use and the disposal of components containing harmful substances, for example, can be traced.

5. Conclusion and Outlook

This article describes the Process Data Twin, embedded within an ecosystem, providing a simple and cross-manufacturer networking solution. The various aspects, both technical and structural, have been discussed and illustrated through use cases. The implementation is carried out within the Industry Fusion Foundation (IFF), which comprises multiple companies collaborating on the solution. In the future, a clear identification as a service solution is intended to be available, serving as an enabler for various use cases, such as Equipment-as-a-Service. Further research aspects arise regarding the active control of machines and topics such as a dynamic safety structure for the utilized machinery and systems.

References

[1] Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In ASME 2017 12th International Manufacturing Science and Engineering Conference (pp. V002T03A003-V002T03A003). American Society of Mechanical Engineers.

[2] Y Tao, F., Zhang, H., Liu, A., & Nee, A. Y. (2018). Digital twin in industry: State-of-the-art. IEEE Transactions on Industrial Informatics, 15(4), 2405-2415.

[3] Kritzinger,W.; Karner, M.;Traar, G. et al.: DigitalTwin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine 51 (2018) 11, S. 1016–1022

[4] Schumacher, A., Eitel, A., & Verl, A. (2018). Administration Shell Concept for Industrie 4.0 Components. Procedia Cirp, 72, 577-582.

[5] IDTA: Specification of the Asset Administration Shell Part 1: Metamodel, Industrial Digital Twin Association, Stand: 2023, Internet: https://industrialdigitaltwin.org/wp-content/uploads/2023/06/IDTA-01001-3-0_SpecificationAssetAdministrationShell_Part1_Metamodel.pdf, 15.03.2024

[6] IDTA: Specification of the Asset Administration Shell Part 3a: Data Specification – IEC 61360, Industrial Digital Twin Association, Stand: 2023, Internet: https://industrialdigitaltwin.org/wp-content/uploads/2023/04/IDTA-01003-a-3-0_SpecificationAssetAdministrationShell_Part3a_DataSpecification_IEC61360.pdf, 15.03.2024

[7] Mahnke, Wolfgang; Leitner, Stefan-Helmut; Damm, Matthias. OPC UA: The Basics. Springer Vieweg Verlag, 1. Auflage, 2009. ISBN: 978-3-8348-0618-2

[8] IEC 62541: OPC Unified Architecture (OPC UA). 2nd edition. Geneva: International Electrotechnical Commission, 2015.

[9] Käbisch, S., et al. (2019). Towards Seamless Data Exchange with JSON-LD. In Proceedings of the 12th International Conference on Semantic Systems (SEMANTiCS '19), ACM, New York, NY, USA, Article 3, 10 pages.

[10] ETSI GS CIM 009 V3.2.1 (2020-07): Next Generation Service Interface (NGSI) for smart cities; Part 2: Interface specification for context management.

[11] Knublauch, H., Kontokostas, D., Hellmann, S., Lehmann, J., & Auer, S. (2017). SHACL: A specification language for validating RDF data. Semantic Web, 8(3), 429-443.

[12] 1. W3C (World Wide Web Consortium). (2017). SHACL: Shapes Constraint Language. Verfügbar unter: https://www.w3.org/TR/shacl/, 11.03.2024

[13] Link for the IndustryFusion GitHub: https://github.com/IndustryFusion/DigitalTwin

[14] Kegel, J., & Jasperneite, J. (2017). Plug-and-Produce System Architecture for Flexible Production Systems. In 2017 IEEE 22nd International Conference on Emerging Technologies and Factory Automation (ETFA) (pp. 1-4). DOI: 10.1109/ETFA.2017.8247756.

[15] www.industry-fusion.org, 20.06.2024

Biography

Andreas Leinenbach (*1991) has been a member of the Chair of Digital Manufacturing at the University of Augsburg since 2023. His research interests include the intelligent production automation.

Marcel Wagner (*1971) is a Principal Engineer in the Network and Edge Group of Intel. His research interests include real-time analytics, semantic data processing and Cloud/Edge architectures.

Fabio Oettl (*1994) currently works as Managing Director of the IndustryFusion Foundation. Prior to this position, he was a research assistant at the Chair of Digital Manufacturing at the University of Augsburg from 2019 to 2023. His main research focus was the identification and evaluation of the application fields of the digital twin and their introduction in companies.

Johannes Schilp (*1976) has been Chair of Digital Manufacturing since 2015. He is also head of the processing department at the Fraunhofer Research Institution for Casting, Composite and Processing Technology and a member of scientific networks like WG MHI and automation research groups (VDMA).