Grant Agreement No. 767498
Innovation Action Project
H2020-FOF-12-2017

D4.1.

Functional and Modular Architecture of MIDIH CPS/IOT System v1

WP4 - Open Platform architecture, development, integration and testing

PUBLIC VERSION

Table of Contents

List of Figures

Executive summary

The D4.1 "Functional and Modular Architecture of MIDIH CPS/IOT System v1" is a confidential document delivered in the context of WP4, Task WP4.1: Functional and Modular Architecture, Open Platforms. Task WP4.1 aims at the definition of a functional and modular architecture that supports IoT, Big Data and Artificial Intelligence technologies, which are expected to drive the change in Manufacturing Industry by enabling smart products (digital inside), smart processes and smart business models.

Based on the description of a Reference Architecture specified in the International Standard *ISO/IEC/IEEE 42010 Systems and software engineering — Architecture description*

> *A Reference Architecture **describes the structure of a system with its elements types and their structures, as well as their interaction types, among each other** and with their environment. Describing this, a Reference Architecture defines restrictions for an instantiation (concrete architecture). Through abstraction from individual details, a Reference Architecture is universally valid within a specific domain. Further architectures with the same functional requirements can be constructed based on the reference architecture. [ISO/IEC42010]*

A reference architecture is a reference model capturing the main architectural characteristics of a set of systems in the same domain, that is mapped onto the software elements that implement the functionality defined in the model, providing indications and guidelines for the design of architectures for a specific system. Therefore, the design of the MIDIH Reference Architecture aimed to identify and define at high-level the functionalities that can enable all the cross-border experiments covered in the CPS/IOT domain, and consequently, to identify and structure through different functional layers a set of logical components to cover these functionalities and their interactions.

As a starting point of the design of the MIDIH RA, a detailed review and analysis of the reference models and architectures for IIoT systems such as IIRA, IVRA and RAMI4.0, as well as an analysis of IoT standard technologies applied in manufacturing domain have been made. In the same way, relevant projects developed in some European initiatives providing Digital Industrial Platforms, have been identified (BEinCPPS, ArrowHead, FAR-EDGE, C2Net and SymbIoTe) and the logical layers that made up their Reference Architectures have been reviewed. This high-level analysis has mainly focused on identify commonalities among them, both at the level of functionalities and technologies, covering standards, interoperability, connectivity, analytics, security and mechanisms to access information.

The document describes the MIDIH RA based on FIWARE for Industry solution, but following a data-driven approach, empowered by Industrial IOT and Industrial Analytics new functions and by specific solutions to interoperate with the Factory Automation real world and with the brownfield of legacy and proprietary systems, which will allow industries making strategic decisions based on data analysis and interpretation in real or near real-time. For this purpose, the core of the MIDIH RA will cover a unified analytics framework interconnected contemplating the Data in Motion (Industrial IOT) and the Data at Rest (Industrial Analytics).

Access and authentication will play a relevant role in the MIDIH RA, based on rules ensuring that only the right person gets the right data at the right time.

Finally, the document proposes two different lines to implement the MIDIH RA (based on FIWARE4Industry and APACHE), identifying for it several components as possible candidate to implement the logical architectural components and satisfy their functionalities. This initial selection will be refined for the second iteration of the D4.1.

This document has been elaborated by the main partners involved in the definition, development, integration and testing of the MIDIH Open Platform architecture (WP4), assisted by the experiment leaders (WP5), either directly or indirectly through the specification of functional components and refining the functional layers of the MIDIH RA.

1 FIWARE4Industry Reference Architecture

FIWARE for Industry (F4I) is a multi-project initiative aiming at developing an ecosystem of FIWARE-enabled software components, suitable to meet the challenges of Manufacturing Industry business scenarios, as indicated by Industry 4.0 vision. F4I originates at the end of 2015 as the exploitation booster of the FITMAN FP7 FI PPP project (www.fiware4industry.com) which developed Open Source reference implementations of Smart-Digital-Virtual Factory scenarios by integrating 14 FIWARE Generic Enablers with 15 original Manufacturing Industry Specific Enablers.

Starting from these 29 components, more than 15 new projects are developing their Smart Manufacturing solutions in several R&I domains. In the Factories of the Future H2020 cPPP, the recent projects CREMA, C2NET, PSYMBIOSYS, BEinCPPS, some of the ongoing FoF11 Connected Factories projects (e.g. FAR EDGE, AUTOWARE, vFOS and NIMBLE) RIAs and two I4MS Phase III Innovation Actions (MIDIH in the domain of CPS/IOT, L4MS in the domain of Mobile Robotics and Industrial Shop-floors' Logistics) are contributing in kind to the picture here below which includes 17 Enablers: 8 enhancements of FITMAN SEs, 3 enhancements of FIWARE GEs, 2 new specific enablers and 4 new releases of FIWARE GEs. Moreover, the EIT DIGITAL High Impact Initiative called OEDIPUS (Operate European Digital Industry with Products and Services) is developing platforms and components FIWARE based for the Smart Manufacturing Industry, in close collaboration with SIEMENS (OEDIPUS coordinator) and its platforms (e.g. MINDSPHERE). In the near future, the EU-Brasil FASTEN project and some National / Regional projects will give their contributions as well to the F4I ecosystem.
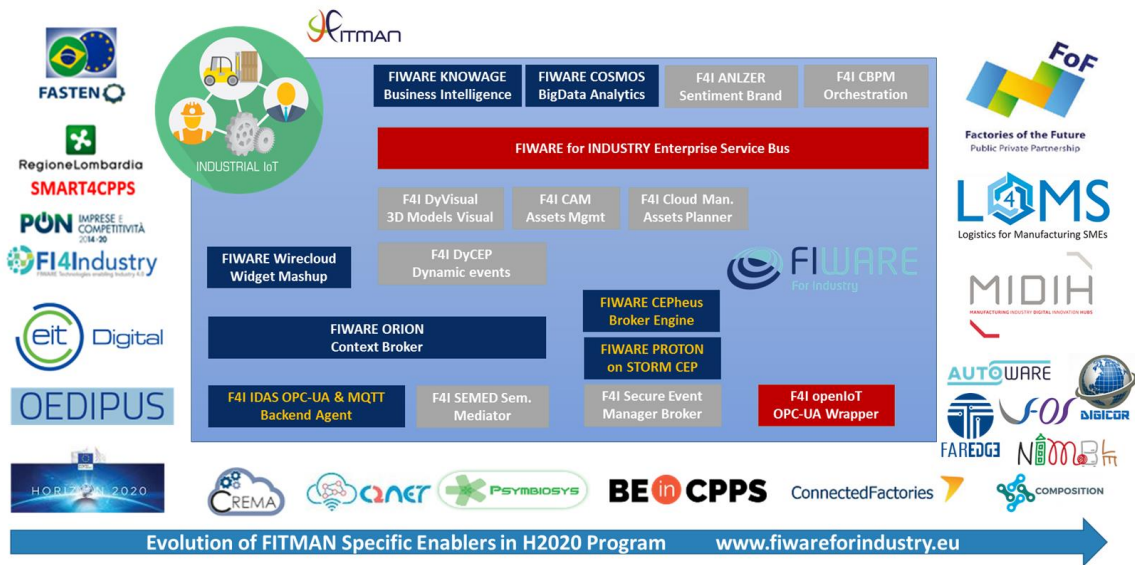
Figure 1. FIWARE4Industry Ecosystem

On the basis of F4I bottom-up success in many projects, the FIWARE Foundation Smart Manufacturing Domain Committee decided to develop a Reference Architecture which could be the common top-down basis for building successful applications on top of both FIWARE GEs and F4I SEs. Following a data-driven approach, the F4I RA is based on the Data in Motion (Industrial IOT) and the Data at Rest (Industrial Analytics) processing modality, what basic components they envisage and how to interconnect them (Lambda architecture) in a unique RA.

2 MIDIH RA for Smart Factory, Smart Product and Smart Supply Chain scenarios

Driven by the requirements of our Industrial Experiments (CRF NECO and IDSA) in the domains of Smart Factory, Smart Product and Smart Supply Chain, MIDIH has elaborated its RA, extending and complementing the F4I RA.

In the domains of Smart Factory and Smart Product we have in fact identified the Real World assets at the bottom of the RA and the Application Ecosystem at the top of the RA. **Data in Motion** is in fact generated by different Real World assets being them in the Smart Factory (Machine Tools, Robots, Warehouses, Human Workspaces) or in the Smart Product (Fleet of Vehicles, Product Service Systems, Smart Objects). **Data at Rest** need to be processed in order to feed AI-based advanced applications in the field of Factory and Product condition monitoring, health diagnosis, predictive maintenance, zero defect quality, energy and waste management, human-machine interaction, Virtual and Augmented reality.
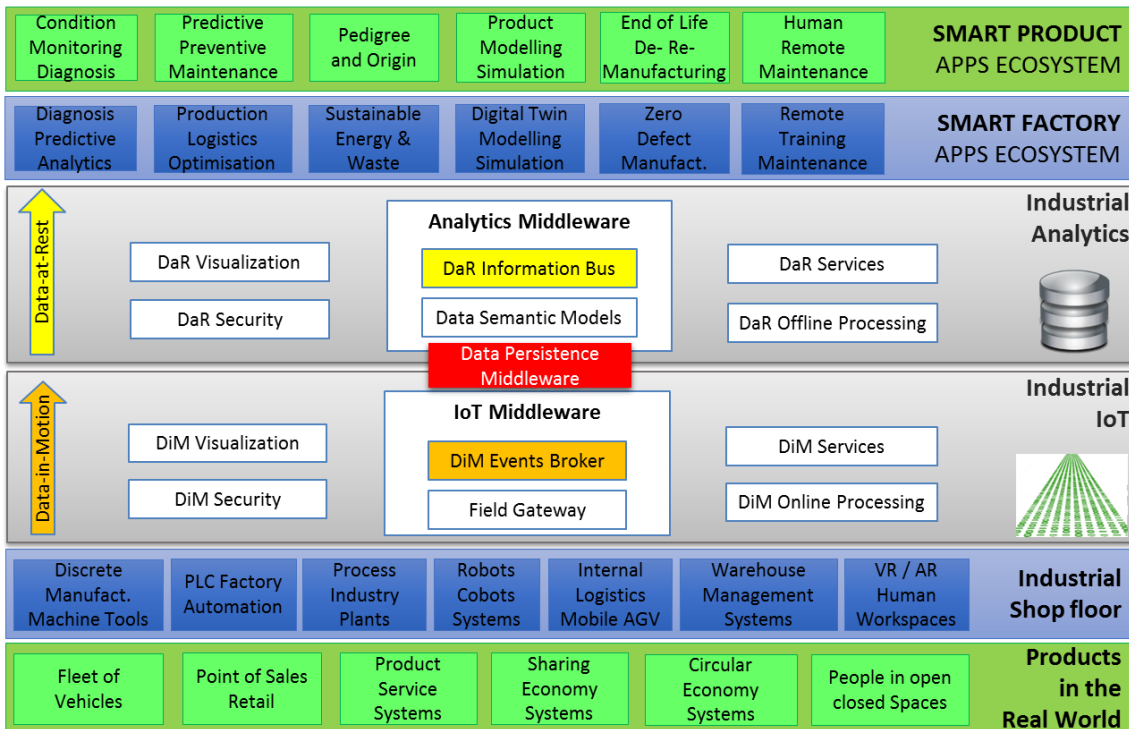
Figure 2. MIDIH RA for Smart Factory and Smart Product

By adopting the IIRA Layered Databus Architectural Pattern and its Machine-Unit-Site and Intersite layers, we've developed the following Data in Motion and Data at Rest model, as a reference for MIDIH Smart Factory and Smart Product industrial scenarios to be implemented in WP5 industrial cases.
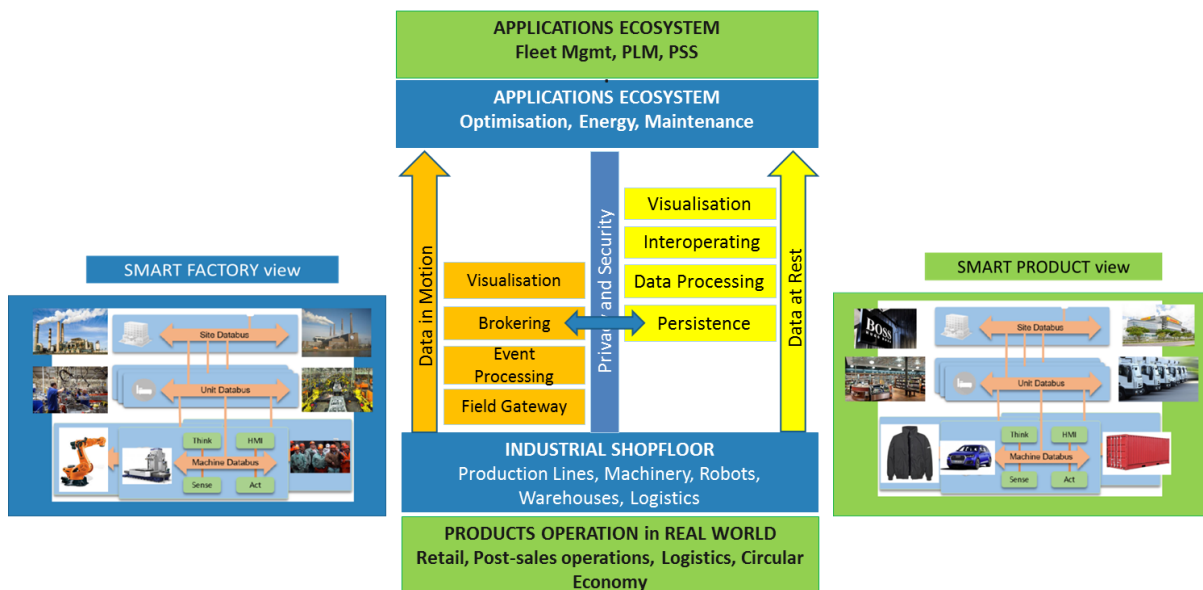

Figure 3. Data in Motion and Data at Rest Model

By leveraging on the mapping between FIWARE GEs, F4I SEs and some further Open Source DiM and DaR components, we have finally compiled an open source reference implementation, which will drive the developments of MIDIH about Data Analytics. This model has two basic lanes, one FIWARE-based and the other based on worldwide known Open Source projects and Foundations, such as APACHE. The main challenge will be to find an interoperability model and implementation between the two Lanes.
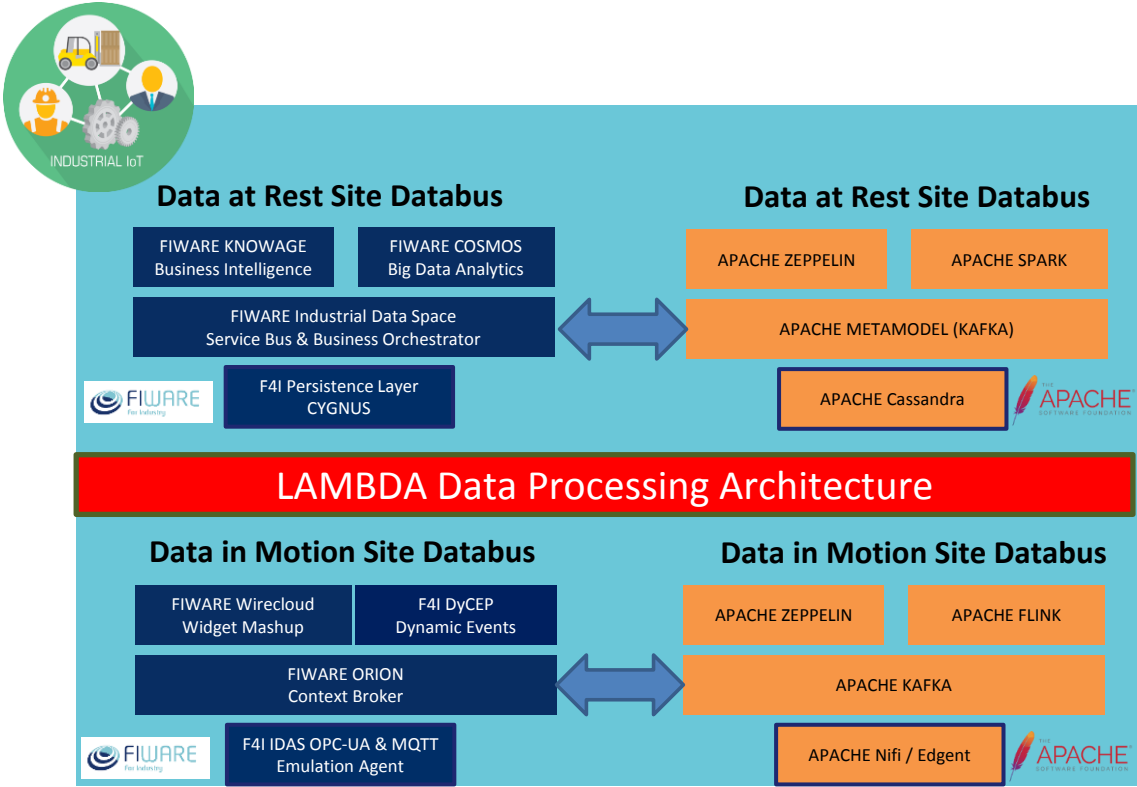

Figure 4. MIDIH Open Source Reference Implementation

In the domain of **Smart Supply Chain**, MIDIH will leverage on the existing open source FIWARE implementation of the IDS RA and extend it with a Distributed Ledger layer to trace and control B2B transactions along the value chains, as required by some of our MIDIH Industrial Cases.
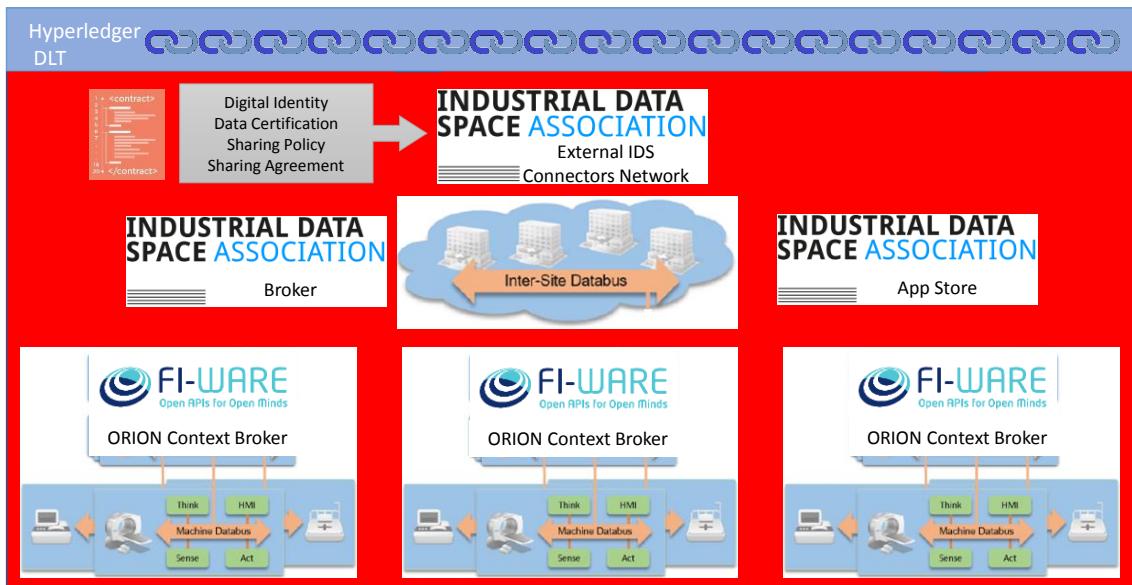
Figure 5. MIDIH RA for Smart Supply Chain

3 MIDIH Modular and Functional Architecture

Considering the RAMI4.0, the IIC reference architecture model (IIRA) and starting on the premise that the MIDIH Reference Architecture will include a "platform" or "middleware" which collects data from multiple sources, stores it, makes it available to various applications such as for analytics, display, and reporting etc. And having special focus on how manage the data-in-motion and the data-at-rest to have a data-driven architectural approach, the objective has been pick the Layered Databus pattern which is an specific instance of the three-tier architecture pattern, and go through the different layers bottom-up identifying functional components or building blocks to later map them to technology that covers these functionalities. In parallel security along the different layers will be taken into account. The requirements coming from the different experiments specified in the three different scenarios: Smart Factory, Smart product and Smart Supply chain have contributed to define some of the functional components that will made up the open and modular MIDIH Reference Architecture.

Leaving the lower layer defined in the *layered databus pattern* which comprises the machine databus, out of the scope of the project, the mapping of the other defined databuses with the three-tier architecture pattern (see *Figure 6*) is as follows:

- The *MIDIH Field or Edge Tier* is mapped with the Unit Databus.
- The *MIDIH Platform and Enterprise Tiers* are mapped with the Site Databus.
- The *MIDIH Business Ecosystem Tier* is a new layer entrusted to the three-tier architecture pattern to enable inter-factory communications that will allow the creation of a reliable MIDIH business ecosystem, and it is mapped with the Inter-Site Databus.
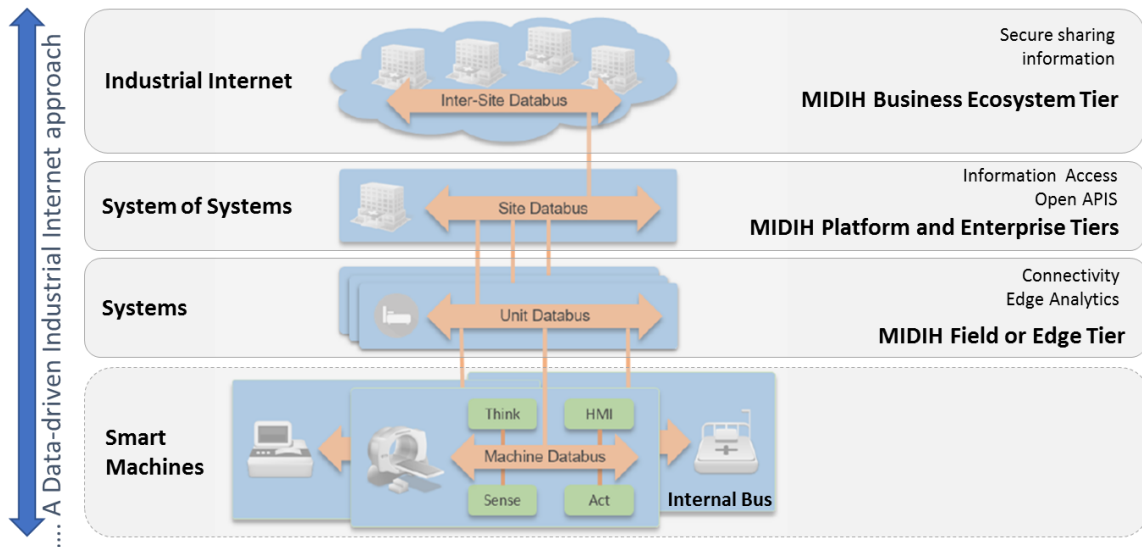
Figure 6. Databuses of layered databus pattern mapping to the Three-tier Architecture

3.1 MIDIH Field or Edge Tier

The Field or Edge tier is mainly used to gather the sensor and machine data from the deployed machines (the sources of data) using various connection types. The edge tier contains intelligent devices and special and general-purpose computers. Edge computing enables some data management and analysis functions to be performed in this tier, in small datasets, using data, applications, and services contained in the edge. Therefore, deciding whether all the data collected is forwarded to the platform tier or data aggregation or processing is needed. By performing small-scale pre-processing and analytics in the edge, where fast decision making is paramount, larger-scale or big data processing in higher tiers becomes easier to manage.

Therefore, the MIDIH edge tier must be modular in its support to multiple protocols for data ingestion, allowing customization of existing and development of new means of communicating with connected assets or devices. Then, it will support a wide ecosystem of heterogeneous devices, with the objective of gather machine data and make it available to the upper layers and IoT systems. In the same way, it must cover security aspects, providing mechanisms for encryption, authentication and data protection functionalities to address elevated enterprise security requirements of connected mission-critical hardware.

3.2 MIDIH Platform Tier

Every industrial company has hundreds of event streams continuously flowing through its networks, conveying sensor data, web clickstreams, copies of business transactions, geolocation data, tweets, market data, weather data, social media activity and other data. The number and volume of streams are growing quickly, driven by the Internet of Things and other market forces. Leveraging the processing of this data enables making smarter, and sometimes faster, business decisions, which is the task of the Platform tier.

According to IIRA, the platform tier receives the data, and ingests and organizes it before storing it in the appropriate data store, and also receives, processes and forwards control

commands from the enterprise tier to the edge tier. It consolidates processes and analyzes data flows from the edge tier and other tiers. It provides management functions for devices and assets. It also offers non-domain specific services such as data query and analytics.

Components from all functional domains may leverage the same data and use analytic platforms and services to transform data into information for their specific purposes.

The main role of this tier is to enable different types of information processing since business intelligence (BI) and analytics modernization programs are undertaking an increasing number of projects that deal with streaming data.

This tier should enable the usage multiple kinds of products to support the diversity of stream analytics applications. Such products include event stream processing (ESP) platforms, stream analytics platforms, business activity monitoring (BAM) platforms, operations intelligence platforms, smart data discovery products and data science platforms.

3.3   MIDIH Enterprise Tier

The MIDIH Enterprise Tier receives the data and integrates it with data from other systems, to perform analysis across business silos, carrying out industry domain-specific business applications, related decision support and business intelligence systems. This tier provides interfaces to human consumers of the information, such as business end users, operators like field service technicians, or the monitoring and diagnostics operation specialists. The enterprise tier may often receive the data flows from the edge and platform tier. At the same time, this tier may also execute control commands on the edge or platform tier.

Therefore, this layer deals with the information architecture, supporting information integration and services capability, and enabling a virtualized information data layer assuring the quality of the data and information.

3.4   MIDIH Business Ecosystem Tier

The MIDIH Business Ecosystem Tier is devoted to enable inter-factory communications that will allow the creation of a reliable MIDIH business ecosystem. Therefore, the main objective of this tier is to provide the mechanisms that allow the exchange of data between different stakeholders within a secure infrastructure, using shared reference architecture and common governance rules covering ownership, access and usage of the data.

3.5   MIDIH Modular and Functional Architecture

With the objective of having a data-driven architecture and starting from the draft of the MIDIH architecture reported in the Description of Work (*Figure 7*), which extends the Open Source BEinCPPS Platform with new components and functionalities mainly focused on support edge-oriented factory automation architectures including Local Clouds and IEC61499 standard according to the ARTEMIS Arrowhead project; brownfield integration with legacy systems and proprietary platforms in the domain of IOT, PLM, ERP through open APIS, and the integration of state-of-the-art Open Source components from worldwide spread communities especially in the field of Industrial analytics, Enterprise Application Integration and Business

Intelligence; together with the alignment with the relevant RA for IIoT existing nowadays, will be the basis for characterize the open and modular MIDIH RA.
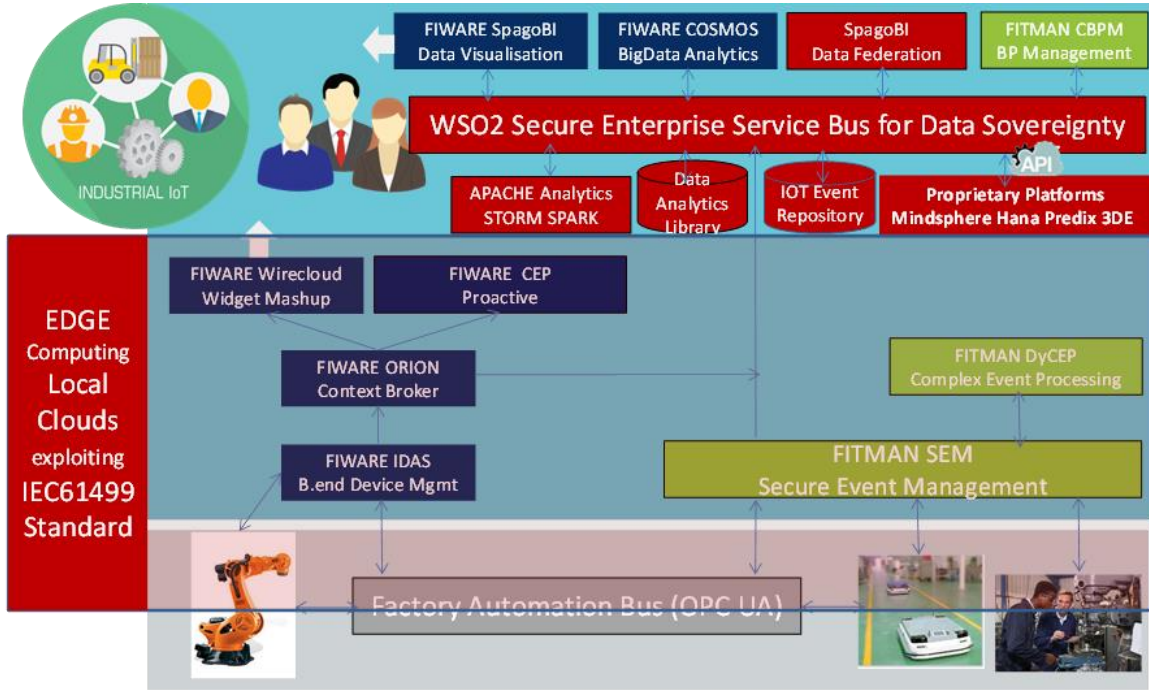

Figure 7. Draft of the MIDIH architecture

Therefore, in the implementation of the MIDIH RA, the Arrowhead distributed automation platform, the FIWARE Data in Motion and Data at Rest platforms and the IDS RA open source implementation via Blockchain Smart Contracts will be integrated into a unique open and modular architecture for Smart factory, Smart Products and Smart Supply Chains industrial scenarios. This architecture will allow integrating and providing solutions in the field of software components, data value chains and innovation business models, giving support to the three main advanced technologies which are expected to drive the change in Manufacturing Industry: IoT, Big Data and Artificial Intelligence. All of this, with some additional open source solutions in order to allow enough flexibility and agility to developers, system integrators and users and avoiding being lock-in to technologies or proprietary solutions.

The final MIDIH RA follows a data-driven approach, based on the Layered Databus pattern defined by IIRA, which foresees four main Databus layers where data exchange / sharing processes are taking place. The Machine Databus is the real world databus, to be implemented in a factory shopfloor, in the operations of a product or in transportation-logistics scenarios and which is outside of the scope of the MIDIH project. The Unit Databus is usually implemented by dedicated edge-fog data gateways as a bridge between Real and Digital worlds. The Site Databus implements the databus of a single administrative domain, being it a company, an IT department, a plant, a fleet of logistics vehicles. On top of that, the Inter-site Databus materializes B2B data exchange / sharing business processes across at least two different administrative domains.

The implementation of the reference architectures in the three scenarios described above shows that Smart factory and Smart Product scenarios are implemented by the first three Databus layers, while the Supply Chain one by the top Intersite Databus.
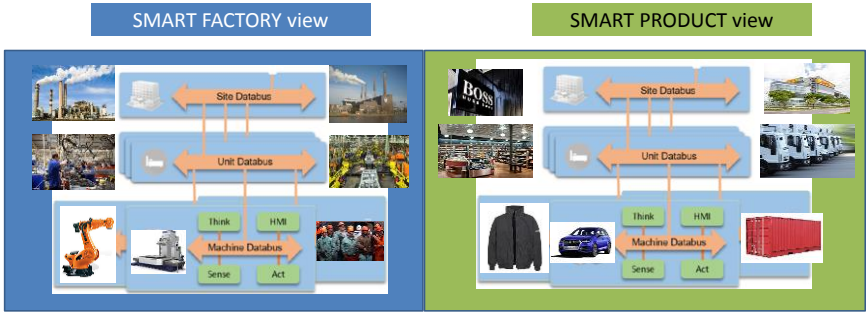
Figure 8. MIDIH Smart Factory, Smart Product and Smart Supply Chain implementation overview

Thus, the Unit Databus will be covered by Arrowhead Data-bus, the Site Databus is instead to be implemented by the FIWARE Orion Context Broker, which is the aggregation point of the FIWARE IoT architecture (data in Motion) and the major Publish /Subscribe broker for FIWARE Analytics applications (data at Rest). And finally, the Inter-Site Databus will be implemented by the FIWARE implementation of the IDS RA extended with a Distributed Ledger layer to trace and control B2B transactions along the value chains. *Figure 9* shows graphically the different databuses and their implementation.

For the implementation of the data analytics framework of the MIDIH RA and with the aim of offering greater flexibility and interoperability, in addition to the implementation based on FIWARE4Industry components, a parallel implementation line based on open source Apache components has been considered. Referring to the different databuses reflected in the previous paragraph, the Apache Kafka which is a high scalable and distributed publish-subscribe messaging system integrating applications/data streams is the selected component to materialize the Unit and Site databuses.
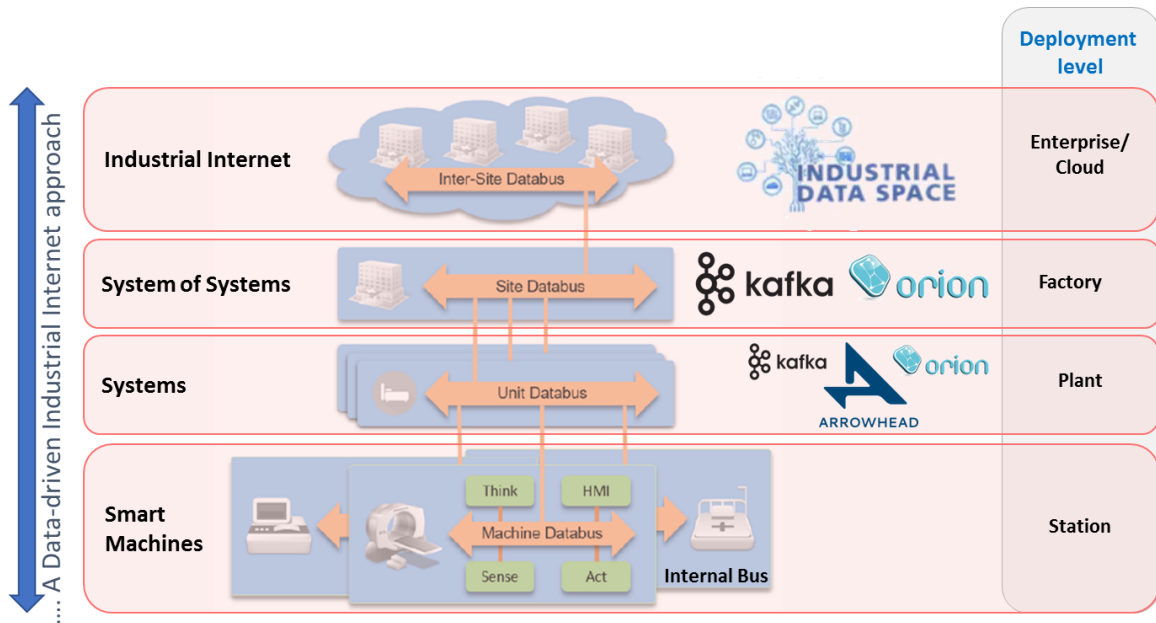
Figure 9. MIDIH Reference Implementation of the Layered Databus pattern

The following picture reflects the core components of the MIDIH data-driven Reference Architecture, focuses on the two layers that made up the MIDIH data analytics framework. The lowest layer is responsible for the data in motion, including the interaction through the Field gateway and specific IoT Agents with multiple devices that use diverse standards and/or protocols, making the gathered data compatible and available to the MIDIH platform. This layer is charge of analysing the streaming data and offering services on these as well as mechanisms for their visualization.

The upper layer is in charge to deal with the Industrial Analytics or data at rest, then responsible to analyse the persisted data either in databases or file systems. It offers as well as the Industrial IoT layer, services that use this data and visualization tools.
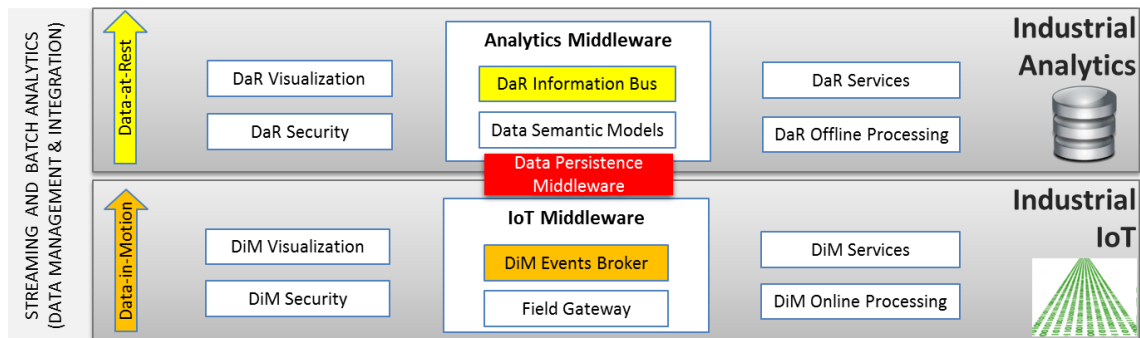

Figure 10. MIDIH Data Analytics Framework for data-in-motion and data-at-rest

Figure 11 shows the different modules on top of the data analytic framework, heart of the data-driven Reference Architecture that made up the final Open and Modular MIDIH RA:

- ***Data Access / Information module,*** this module deals with the information architecture, supporting information integration and services capability, and enabling a virtualized information data layer assuring the quality of the data and information. Therefore, it provides several capabilities of which we can highlight the ability to retrieve data from heterogeneous information sources, transform it into a common format, and expose it to consumers using different protocols and formats, assuring at the same time the security and protection of information. Consequently, it is responsible for basic information management concerns such as metadata and unstructured data management. Furthermore, it can be considered an interoperability point because it is the entry point for external applications to interact with the MIDIH platform in order to access the information.
  To cover all these functionalities, it offers:
    o A set of uniform APIs providing the MIDIH platform functionalities that will be used by the final end-user applications such as business applications and services.
    o Interoperability functionalities, in charge to harmonize the access to the data.
    o Data models, representing the objects and events associations for describing and manipulating data, relationships between data, and constraints on the data in the MIDIH platform.
  For this purpose, The **WSO2[1]**, an Open source Enterprise Service Bus (ESB) can be a valid solution in order to provide integration of enterprise applications and services. It enables diverse applications, services, and systems to talk to each other through a common communication bus, using lightweight and standard messaging protocols such as SOAP and JSON.
- ***Business Apps and services,*** these are the applications and services that have to make use of the data provided by the platform through an unified the data access / information layer
- ***Inter-Site Secure Data-bus,*** this module is responsible to materialize B2B data exchange / sharing business processes across at least two different stakeholders or administrative domains. Then, provides mechanisms to trace and control B2B transactions along the value chains.
- ***Third Party Applications,*** refer to applications provided by vendors either as an open source or as a commercial product and generally specialized in specific domains. Regarding MIDIH RA, it is focused on brownfield integration and interoperability with proprietary solutions and standards in the field of IOT.
- ***Security***, is a transversal module responsible to cover all the security aspects along the MIDIH RA, including Security, Privacy and Governance. This module supports both data, infrastructure, and the services offered. So, main functionalities provided cover authentication, authorization, access control and confidentiality.
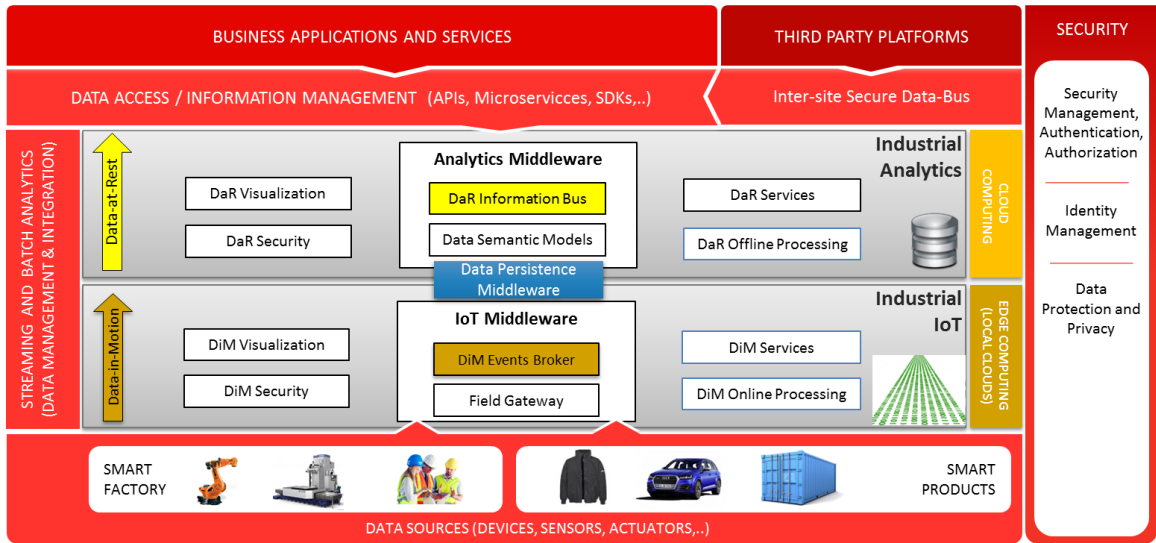
---

[1] https://wso2.com/

Figure 11. Data-driven MIDIH Reference Architecture

The *Figure 12* sketches an envisaged FIWARE and Apache components to cover the functionalities of the MIDIH Reference Architecture. At the moment a first selection has been made that will be refined according to the needs of the use cases.
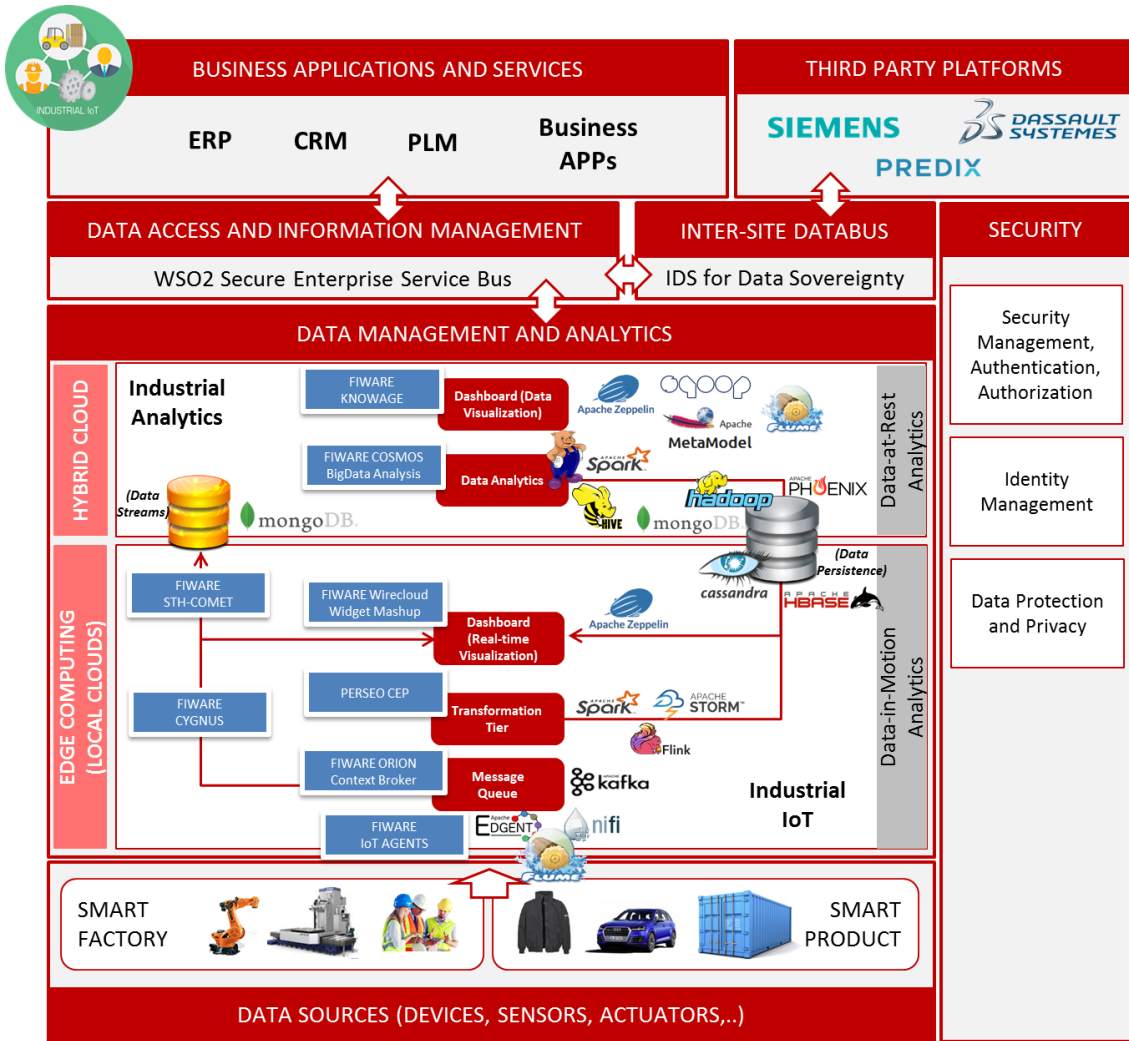
Figure 12. MIDIH RA data analytics framework implementation overview