From Cloud-Native to 5G-Ready Vertical Applications: An Industry 4.0 Use Case

Raffaele Bolla DITEN, University of Genoa, Italy

Franco Davoli DITEN, University of Genoa, Italy S2N National Lab., CNIT, Italy

Chiara Lombardo S2N National Lab., CNIT, Italy Roberto Bruschi DITEN, University of Genoa, Italy

> Zied Ghrairi BIBA, Bremen, Germany

Jane Frances Pajo Telenor Research, Norway Kay Burow BIBA, Bremen, Germany

> Panagiotis Gouvas UBITECH, Greece

Anastasios Zafeiropoulos National Technical University of Athens, Greece

Abstract—This paper aims to showcase the ability of the MATILDA Platform to enable vertical applications fully utilizing the capabilities offered by the fifth generation of mobile networks (5G). Although 5G, powered by network slicing and edge computing, promises to flexibly support radically new and extremely heterogeneous vertical applications, vertical stakeholders generally lack both the skills to exploit the full potentials of 5G networks and the vision of the underlying resources, owned by Telecom providers that are reluctant to expose them in an unmediated way. The MATILDA platform bridges the gap between the vertical application and the network service domains. This paper presents the case of an Industry 4.0 application, and highlights the role played by the MATILDA solution in its successful deployment and orchestration.

Keywords—5G, slicing, Industry 4.0

I. INTRODUCTION

The advent of the fifth generation of mobile networks (5G) is predicted as a leading factor in the digital transformation by Big Data and Cloud Computing technologies towards a new hyper-connected society. 5G has been designed to provide flexible support for radically new and extremely heterogeneous vertical applications, requiring any custom mix of network and radio services. In particular, applications under the Industry 4.0 umbrella, with their zero-perceived latency requirements, would strongly benefit from the 5G widespread diffusion. As market trends show, the demands for new services and new technical solutions coming from stakeholders and customers postulate the industrial sector, which has to react on these demands and further develop the existing manufacturing towards a higher level of flexibility, efficiency and "servitization": on one side, to provide the requested high level of servitization up to services as a service (SaaS), on the other side, to further integrate and adapt highly automated manufacturing systems, which are running on technical and physical limits. Especially the latter show the need for new solutions and approaches [1].

In this context, 5G network slicing [2] and edge computing [3] can provide the required connectivity features by hosting vertical applications onto the 5G infrastructure, and directly attaching them to the network slice terminations in neighboring geographical facilities. Network slices can be customized to perfectly match the desired network and radio services features and constraints of vertical industries to interconnect their devices and terminals and directly manage them.

However, vertical stakeholders generally lack the basic knowledge to exploit the full potentials of 5G networks. Moreover, the Telecom providers owning the physical infrastructure are less than prone to allow third parties independently orchestrating their resources. In such context, the necessary integration between the digital systems that enable vertical services and the network layer remains undefined and represents a big challenge.

In this scenario, the main aim of the MATILDA H2020 European Project [4] has been to bridge the gap between the vertical application and the network service domains with a platform able to *i*) transform a cloud-native application into a 5G-ready one, *ii*) providing the vertical stakeholders with a complete 5G network slice, from the radio access to the core, and *iii*) allowing them to manage the lifecycle of their applications, regardless of the underneath infrastructure, all in a Zero-touch network and Service Management (ZSM) fashion [5]. MATILDA provides verticals with a familiar, cloud-style interface in which, by simply adding their application instantiated over a wide-area infrastructure that they can orchestrate in a secure and isolated fashion.

In order to assess how MATILDA is suited to enable vertical applications fully utilizing the capabilities offered by 5G, this paper presents one of the demonstrators developed within the project and presented during the Project's final demonstration in October 2020, showcasing an Industry 4.0 application, and highlights the role played by the MATILDA solution in proving the application successful.

The paper is organized as follows: Section II presents the MATILDA platform, while the Industry 4.0 use case can be found in Section III and the conclusions are drawn in Section IV.

II. THE MATILDA VISION AND SOLUTION

MATILDA comes up with a novel and holistic approach for tackling the overall lifecycle of applications' design, development, deployment and orchestration in a 5G environment. A set of novel concepts have been introduced, including the design and development of 5G-ready applications – based on cloud-native/microservice development principles [6], the separation of concerns among the orchestration of the developed applications and the required network services that support them, as well as the specification and management of network slices that are application-aware and can lead to optimal application execution.

MATILDA's main targets are i) the vertical application stakeholders, whose developers' work can be substantially eased and enhanced, and ii) the Telecom Providers, which MATILDA can support to satisfy their vertical customers' needs.

A. Lifecycle of a 5G-Ready Application

A summary of the overall lifecycle of an application created with the MATILDA framework is represented in Figure 1, highlighting the interaction among the different stakeholders and the usage of metamodels.

The MATILDA reference architecture is divided in three distinct layers: namely, the 5G-ready Applications Layer, the Applications' Orchestration Layer and the Network and Computing Slice Management Layer. Separation of concerns per layer is a basic principle adhered towards the design of the overall architecture, which facilitates the adoption and integration in their systems by the involved stakeholders, and allows the separate installation of individual parts of the platform, as well.

The 5G-ready Applications Layer is oriented to software developers. It takes into account the design and development of 5G-ready applications per industry vertical, along with the specification of the associated networking requirements, and defines the business functions, as well as the service qualities of the individual application.

The Applications' Orchestration Layer is oriented to application service providers and supports the dynamic onthe-fly deployment and adaptation of the 5G-ready applications to their service requirements, by using a set of optimization schemes and intelligent algorithms to provide the needed resources across the available multi-site programmable infrastructure.

Finally, the Programmable 5G Infrastructure Slicing and Management Layer, oriented to telecommunications infrastructure providers, is responsible for setting up and managing the 5G-ready application deployment and operation over an application-aware network slice. Such operations are triggered based on requests provided by the Applications' Orchestration Layer through the specification of Open Application Programming Interfaces (APIs).

B. The Vertical Industry Domain

The framework allows software developers to create applications following a simple and conventional microservices-based approach, where each component can be independently orchestratable. Based on the conceptualization of metamodels (application component and graph metamodels), they can formally (and easily) declare information and requirements - in the form of descriptors that can be exploited during the deployment and operation over programmable infrastructure. Such information and requirements may regard capabilities, envisaged functionalities and soft or hard constraints that have to be fulfilled and may be associated with an application component or virtual link interconnecting two components within an application graph. The produced application is considered to be "5G-ready".

MATILDA also encourages new business and wide collaboration by providing a Marketplace, where not only the created applications and components can be published but also Virtual Network Functions (VNFs) and Network Services (NSs) (in the form of enhanced descriptors).

Application Service Providers are able to adopt the developed 5G-ready applications (published to the Marketplace or created internally) and specify policies and configuration options for their optimal deployment and operation over programmable infrastructure. Based on the provided application descriptor, these service providers are able to design operational policies and formulate a slice intent. This information is used by the Vertical Application Orchestrator (VAO) to request the creation of an appropriate application-aware network slice from the Telecommunication Infrastructure Provider.

While the instantiation and management of the application-aware network slice (including the set of network

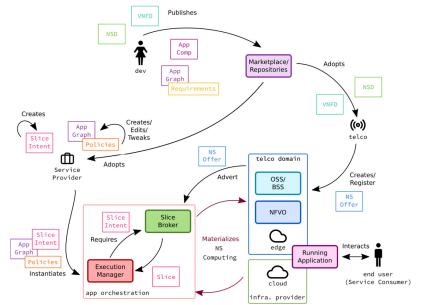


Figure 1. MATILDA workflow highlighting the different stakeholders and metamodels.

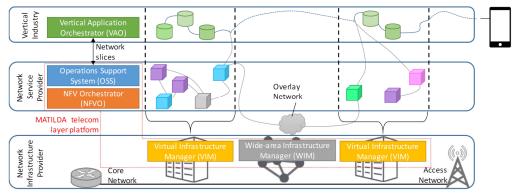


Figure 2. Example of deployment of an application, driven by the MATILDA framework, into multiple VIMs over a 5G infrastructure.

functions) is realized by the Network and Computing Slice Deployment Platform, the deployment and runtime management of an application is independently realized by the VAO following a service-mesh-oriented approach. Advanced monitoring and analysis techniques are also applied for extracting insights that can be proven useful for service providers.

C. The Telco Domain

Telecommunication Infrastructure Providers rely on the concept of network slice to fulfil the vertical applications' needs during the overall lifecycle of the 5G-ready application. A network slice is a logical infrastructure partitioning allocated resources and optimized topology with appropriate isolation, to serve the particular purpose of an application graph. The Network and Computing Slice Deployment Platform in Figure 2 includes an OSS/BSS system, a Network Functions Virtualization Orchestrator (NFVO), a Wide Area Infrastructure Manager (WIM) and Virtual Infrastructure Managers (VIMs). It is worth noting that all of these building blocks and their reference points are fully compliant with the specifications of the ETSI NFV [7] architectural framework.

Based on the interpretation of the provided slice intent, the Telecommunication Infrastructure Provider is in charge of realizing the instantiation of the slice over the programmable infrastructure by activating and dynamically handling the required network management mechanisms. The reserved resources for this slice combine both network and compute ones.

The materialization of the network slice requires the instantiation of NSs that can be imported into the telecommunications infrastructure provider's catalogue from the MATILDA marketplace.

Following the generation of the slice intent according to the specifications provided by the application developer, the OSS controls (in a ZSM, automated way) all the resources/services at any layer in the Telco domain that allow the entire lifecycle management of the 5G-enabled application (5G vApp): from planning, to first deployment, down to inlife management (the so-called Day-0, Day-1 and Day-2 operations, respectively), until their termination. Day-2 operations include, among others, upgrade and scaling mechanisms. These automated operations are common to all instantiated vertical applications. Details on the Network and Computing Slice Deployment Platform can be found in [8].

III. HUMAN-ROBOT COLLABORATION IN ASSEMBLY LINES

With the development of Industry 4.0, the automation processes in industrial environments increase, posing important requirements for safe collaboration between human beings and machines in the workspace environments. The manufacturing industry has begun to develop the trend of customization, individuation, and flexibility. One very challenging application of manufacturing and automatization as part of Industry 4.0 is leading to close contact between robots and workers, while the traditional industrial robot production methods cannot meet the needs of the production safety. Therefore, for human-machine collaboration, security aspects have become a top priority [9].

This section presents one of the demonstrators showcased during the final MATILDA review in October 2020. In order to minimize the network latency for the stream, proximity is required among the IP Camera and the Stream Server. For this reason, the whole deployment had taken place in Athens, with the VAO setting the locality constraint through the slice intent.

The remaining of this section analyses the main requirements to be taken into account (Section III.A), the description of the use case (Section III.B), and the benefits of employing the MATILDA platform are highlighted in Section III.C.

A. Main Industry 4.0 Requirements

Industrial end users must be provided with easy-tointegrate functional-safe systems that can enable a higher collaboration level in terms of varying dynamic safety distance between humans and industrial robots, ensuring both human safety and flexible use in existing and future assembly lines, and that are largely independent of specific robot types. In particular, in a Human-Robot Collaboration (HRC) scenario, a significant challenge is to transform an automated process into a collaborative set of process tasks and to protect the human operators in a common/ shared workplace, mainly by co-implementing the following two key safety aspects:

- detection of possible collisions in real time
- adaptive machinery control for run-time human collision avoidance.

For this purpose, the implementation is necessary of varying dynamic Speed and Separation Monitoring (SSM) safety measures based on time efficient approaches of measuring human-robot separation distance and relative speeds.

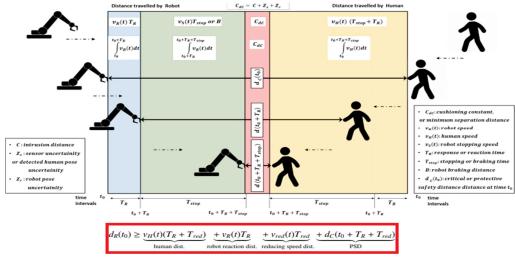


Figure 3. A time interval diagram representing the SSM formulation for defining the critical/protective safety distance between a human and a robot at time t₀ [10].

The key requirement here, which is resulting in a safe and higher collaboration, is to reduce the safety distance between human and robot, by minimizing the end-to-end latency between the human and robot movements' detection and the reaction time of the machinery [10]. As depicted in Figure 3, the calculation of the safety distance is time-based. Taking into consideration the aforementioned aspects, this can be further broken down to the optimization of both the monitoring mechanism (e.g., human detection algorithm) and the feedback control process. To optimize the detection algorithm, one has to focus on minimizing the amount of time required to output a valid decision, while feedback control process optimization involves quick transmission, routing, and dispatching of the produced feedback through the underlying network.

The overall implementation concept relies on direct and continuous position monitoring of the human operator and robot system extremities in the collaborative area.

When a human presence is detected within the coverage area of the robot system, depending on the speed and direction of human operators and robot, the adequate decision has to be taken. This consists of controlling the robot system in an adaptive way depending on the distance between the human and the robot. For this application, the monitoring of the collaborative area can be performed by the use of various indoor positioning technologies like Infrared (IR), Radio Frequency (RF), Pedestrian Dead Reckoning (PDR)/INS and Hybrid.

B. The Industry 4.0 Use Case

The Industry 4.0 application chosen as one of the MATILDA demonstrators involves a human moving towards

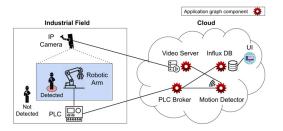


Figure 4. Industrial Production Scenario for the Robotic Arm Control.

or close to the working space of the robotic arm and targets the safety concept presented above. The monitoring of the human operator and robot system is done through a video stream from a camera located on top of the monitored machinery. For this reason, a physical test scenario was developed, which will represent the industrial environment in a simplified manner to make it possible to integrate the set-up into the MATILDA framework. This also means that the hard requirements for an industrial application, especially for HRC, had to be adapted and soften up (e.g. reliability, latency) as the actual physical use case was designed on existing 4G technologies.

The scenario is depicted in Figure 4, along with the cloudnative application components that enable its automation. The involved physical devices are the robotic arm, an IP camera to monitor the area around it, and a Programmable Logic Controller (PLC) that implements the control actions upon the arm. The cloud application is composed by: i) a video streaming server; ii) a motion detector function for real-time distance calculation; iii) a broker that publishes meaningful events regarding identified objects to the PLC; iv) a database for storing the produced events ("Influx DB") along with a data visualization user interface that visualizes the relative positions of the identified objects around the robotic arm (shown as "UI" in Figure 4).

The IP camera (first component) inspects a designated area around the robotic arm and streams the H.264/AVC-encoded video [11] toward the video stream server, which is the first component of the vertical cloud application. Its sole purpose is to transcode the received video frames in order to produce a lightweight equivalent of the video stream.

The transcoded video frames are then transmitted to a motion detector (i.e., the second component of the vertical application graph). This component performs two operations: i it renders the video frames and ii it inputs the rendered frames to a motion detection algorithm. This algorithm employs a minimum enclosing circle scheme [12] to calculate the proximity between the robotic arm and an identified object.

Each of the detected objects is represented as a circle, created with a radius proportional to its speed, along with a set of coordinates that are used to devise its physical distance

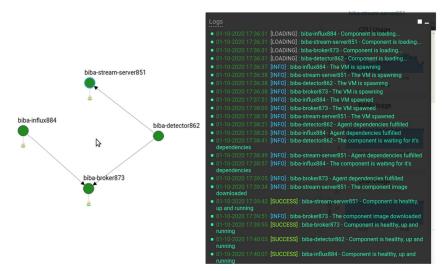


Figure 5. Post-deployment application graph.

from the robotic arm. These coordinates are sent to the PLC broker (i.e., the third application graph component) and are in turn communicated to the PLC device through a secure connection channel. Having those coordinates and knowing the exact location of the robotic arm, the PLC knows whether the robotic arm is likely to collide with any of the detected object(s) or not, and thus can act accordingly (e.g., halt motion, reduce speed, change path...). The fourth component of the vertical application graph is the Influx database instance connected with a graphical user interface for real-time data visualization.

The post-deployment application graph is shown in Figure 5, along with the log of the automated operations triggered by the orchestrator. In practice, what the developer needs to do to create the application graph and trigger its automated deployment is to fill in the forms that specify the characteristics of the application to be conveyed to the VAO,

including the hard or soft constraints requested for, regarding the communication needs of the application components.

The MATILDA project has closely monitored the work performed by standard development organizations - mainly 3GPP, ETSI and ITU - as well as industry alliances and forums. In particular, ITU initially provided the applications and network services generic target Key Performance Indicators (KPIs) that have been refined throughout the project's lifetime. A summary of the collected operational and network KPIs is reported in Table 1. We have run around a thousand runs on the platform deployment described in [8], the confidence interval is 95%. We have activated IEEE Precision Time Protocol [13] to align the clocks of the devices involved in the test runs. KPIs have been collected via software probes in the absence of channel interferences.

It is worth noting that the table does not report the latency figure: the performance evaluation had to be carried out by

Operational KPI	Description	Acceptance Threshold	Result
Resource Usage	Compute/storage/networking resource usage	Must be available	Available.
Monitoring	monitoring		
Scaling time	Time required to trigger the scaling after threshold was	~ 30s	7s
	reached		
Availability	Availability of the service upon request	High > 99 %	99.999 %
Reliability	Reliability of an available service	High > 99 %	99.999 %
Network KPI	Description	Acceptance Threshold	Result
Availability	End-to-end continuous measurements of network	> 99 %	99.5 %
	connectivity, collected statistics used to calculate		
	availability		
Reliability	End-to-end continuous measurements of network	>99 %	99.5 %
	connectivity		
Bandwidth	Required for delivering rich video stream to the	~10 Mbps	~ 12 Mbps
	Detector component in order to make proximity		
	analysis		
Jitter	Time-critical communications should be stable and	< 1 ms	0.7 ms
	reliable. Timing variation must be minimal		
Packet Loss	Reliability and high availability of the services in HRC	< 0.01 %	< 0.001 %
	needs to be guaranteed. Therefore, packet loss should		
	be as small as possible.		

Table 1. Measured KPIs for the Industry 4.0 use case.

using 4G technologies with a number of solutions for emulating network slicing capabilities, such as a VNF providing S1 bypass capabilities [14]. The computation capacity required by the system produced a latency slightly above the 100 ms target.

C. Using the MATILDA Platform

A major advantage offered here by the MATILDA platform, with respect to a more "traditional" implementation that would not rely on the MATILDA orchestration features, is the guaranteed reliability and high availability in the interconnection of distributed manufacturing facilities, which is made possible by the realization of a network slice dedicated to serve the application's needs. Similar, recent projects in the field focus on the NFV aspects [15], [16], or on the transition from cloud-native to 5G-ready applications , but none of them, at least during the 3GPP Phase 2, attempted to tackle both issues.

Moreover, real-time information collected and transferred from and to production processes via the MATILDA monitoring tools enables a better access to remote process control and process monitoring. This will not only increase the Quality of Experience (QoE)/Quality of Service (QoS), but also the performance and flexibility in the use case scenarios. The presence of 5G New Radio (NR) wireless networking equipment would be necessary to contribute ensuring the satisfaction of the tight delay requirements of this application, by providing the benefits of Ultra-reliable and/or Low Latency Communications (URLLC) services for the wireless segment. However, the MATILDA support of edge computing also helps ensure the compliance of end-to-end delay requirements, by deploying intensive data processing tasks in the close proximity of the involved humans and machinery. Additionally, the capability of supporting secure and private end-to-end services is highly relevant in business scenarios where data transfer between interconnected manufacturing facilities is necessary.

Besides these advantages, the application deployment and orchestration platform offers tools to the application developer to specify the communication needs and constraints of the microservice components and to deploy the application graph and onboard the application components in very short time (2 and 5 min have been demonstrated, respectively, for these operations). This is accompanied by the automated generation and deployment of the network slice, once the slice intent, based on the developer's specifications, has been passed to the OSS in the network domain.

IV. CONCLUSIONS

While the fifth generation of mobile networks was designed to support radically new and extremely heterogeneous vertical applications, the lack of knowledge on 5G of the vertical stakeholders and the complexity in orchestrating the computing and network resources owned by Telecom providers may hamper the delivery of truly groundbreaking applications.

In this context, the MATILDA Project aimed to tackle the challenging integration between the vertical services and the networking environments. The released ZSM platform supports verticals in converting their cloud applications into 5G-ready ones, providing them with a complete 5G network slice and to manage the lifecycle of their application. In order to showcase such capability, this paper presented an Industry

4.0 application, which was one of the project demonstrators, including a set of KPIs attesting the effectiveness of the solution in fulfilling the stringent requirements demanded by such applications.

ACKNOWLEDGMENT

This work has been partially supported by the Horizon 2020 5G-PPP Innovation Actions MATILDA and 5G-INDUCE (Grant Agreements no. 761898 and no. 101016941, respectively) and by the Horizon 2020 Innovation Action SPIDER (Grant Agreement no. 833685).

REFERENCES

- A. Osseiran, J.F. Monserrat, P. Marsch, 5G Mobile and Wireless Communications Technology, pp. 16-49, Cambridge University Press, 2016.
- [2] K. Samdanis et al., "5G Network Slicing Part 1: Concepts, Principles, and Architectures," IEEE Commun. Mag., vol. 55, no. 5, pp. 70–71, Aug 2017.
- [3] "Mobile Edge Computing (MEC); Framework and Reference Architecture," ETSI GS MEC 003, v. 1.1.1, March 2016.
- [4] MATILDA A Holistic, Innovative Framework for Design, Development and Orchestration of 5G-ready Applications and Network Services over Sliced Programmable Infrastructure, http://www.matilda-5g.eu/.
- [5] M. Jie *et al.*, "Zero-touch network and service management," Deutsche Telekom, Bonn, Germany, White Paper, Dec. 2017.
- [6] P. Jamshidi, C. Pahl, N. C. Mendonça, J. Lewis, S. Tilkov, "Microservices: The Journey So Far and Challenges Ahead," IEEE Software, vol. 35, no. 3, 2018, pp. 24–35.
- [7] ETSI "Network Functions Virtualisation (NFV), Management and Orchestration," ETSI GS NFV-MAN 001 V1.1.1, URL: http://www.etsi.org/deliver/etsi_gs/NFV-MAN/001_099/001/01.01.01_60/gs_nfv-man001v010101p.pdf
- [8] R. Bruschi, F. Davoli, F. Diaz Bravo, C. Lombardo, S. Mangialardi, J. F. Pajo, "Validation of IaaS-based technologies for 5G-ready applications deployment", Proc. European Conf. on Networks and Commun. 2020 (EuCNC 2020), Dubrovnik, Croatia, 2020.
- [9] 5G for Connected Industries and Automation (White Paper -Second Edition), https://www.5g-acia.org/fileadmin/5G-ACIA/Publikationen/Whitepaper_5G_for_Connected_Industries_and _Automation/WP_5G_for_Connected_Industries_and_Automation_D ownload_19.03.19.pdf.
- [10] S. P. Kumar, "Dynamic Speed and Separation Monitoring with On-Robot Ranging Sensor Arrays for Human and Industrial Robot Collaboration," PhD Thesis, Rochester Institute of Technology, Rochester, NY, USA, 2020. Accessed from https://scholarworks.rit.edu/theses/10391.
- [11] ITU-T H.264 (06/2019) Advanced video coding for generic audiovisual services.
- [12] https://www.geeksforgeeks.org/minimum-enclosing-circle-set-2welzls-algorithm/.
- [13] "IEEE Std. 1588-2002", IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems, Nov. 2002.
- [14] R. Bruschi, F. Davoli, G. Lamanna, C. Lombardo, S. Mangialardi, J. F. Pajo, "Enabling Edge Computing Deployment in 4G and Beyond," in Proc. 6th IEEE Int. Conf. on Network Softwarization (NetSoft), Ghent, Belgium, 2020.
- [15] 5G Development and Validation Platform for Global Industry Specific Network Services and Apps. Online: https://www.5gtango.eu/
- [16] 5th Generation End-to-end Network, Experimentation, System Integration, and Showcasing, <u>https://5genesis.eu/</u>
- [17] 5G-MoNArch: 5G Mobile Network Architecture for diverse services, use cases, and applications in 5G and beyond, Online: <u>https://5g-monarch.eu/</u>.
- [18] NGPaaS: Next Generation Platform as a Service, Online: <u>http://ngpaas.eu/</u>