Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of some Toulouse researchers and makes it freely available over the web where possible.

This is an author's version published in: https://oatao.univ-toulouse.fr/19477

Official URL:

To cite this version:


Any correspondence concerning this service should be sent to the repository administrator:

tech-oatao@listes-diff.inp-toulouse.fr
Abstract—The industrial manufacturing sector has substantially evolved over the past decade, with the collect and analyze of sensors and apparatus data. The manufacturing processes “digitalization” has resulted in huge data management concerns, as manufacturing companies need to aggregate value for productivity improvement. The Industry 4.0 concept involves many technologies in order to challenge this issue. The paper presents an architecture framework to allow data analysis in different factories in the Industry 4.0 context. A MBSE approach is proposed and the use of SysML is assessed in that context.

Keywords—System Modeling, SysML, Industry 4.0, Advanced manufacturing, Big Data Analytics.

I. INTRODUCTION

Over the last century, three industrial revolutions have increased productivity [1]: mechanization, mass production, and automation [2]. Since early 2010’s, a fourth industrial revolution has emerged, relying on new technologies: the Internet and embedded systems have allowed emergence of cyber-physical systems (CPS), the Internet of things (IoT), cloud computing/manufacturing, and Big Data/Analytics. These transformations mark the transition from current industrial production to Industry 4.0 (or fourth industrial revolution), which is characterized by smartness and networking [1].

In that context, the project discussed by the paper aims to develop and implement a pilot integrated system for intelligent automated decisions making in the context of Industry 4.0 (advanced industry). The project contextualizes Industry 4.0 technologies in industrial scenarios, such as cloud computing/manufacturing, IoT, Big Data/Analytics, and digitalization. A general framework, including an architecture, allows data analysis in various scenarios and situations. The proposal is sustained by SysML [3] [4] [5] models developed using the free software TTool [6].

The paper is structured as follows. Section 2 surveys related work on current research in advanced manufacturing topics. Section 3 sketches a methodology. Section 4 introduces the SysML tool TTool [7] [6] to support that methodology. The use of SysML is the subject of Section 5. Section 6 presents the architecture of the advanced manufacturing platform, Big Data/Analytics concepts and user interface. Section 7 concludes the paper and outlines future work.

II. RELATED WORK

A. Industry 4.0 Technologies

To stay competitive in an international and competitive context, industries have to challenge the following issues [8]: short development periods, individualization on demand, flexibility, decentralization and resource efficiency. Some technologies of the Industry 4.0, e.g. CPS, cloud computing/manufacturing, Big Data/Analytics and digitalization, are resources enabling these challenges.

Cyber Physical Systems represent a new generation of digital systems with two main functional components [9]: (i) advanced connectivity ensures real-time data acquisition from the physical world and information feedback from the cyber space [9]; (ii) intelligent data management, analytics and computational capability construct the cyber space [10]. Using CPS improves the implementation of large-scale systems by increasing the adaptability, autonomy, efficiency, functionality, reliability, safety, and usability of such systems.

Cloud computing applied in manufacturing context, also called ‘cloud manufacturing’, is a new manufacturing paradigm based on networks. Zhang et al. [11] defined cloud manufacturing as “the transformation of manufacturing resources and manufacturing capabilities into manufacturing services, which can be managed and operated in an intelligent and unified way to enable the full sharing and circulating of manufacturing resources and manufacturing capabilities”. Cloud manufacturing provides safe, reliable, high quality, low-cost and on-demand manufacturing services for the whole life cycle of manufacturing [11].

According to Lee et al. [12] “big data analytics is a set of processes for retrieving the correct data from high volume, high velocity, and high variety data; identifying patterns in the
data; and improving business decisions based on the results”. Since big data analytics have been identified as a key issue to achieve innovations in manufacturing, many industrial companies have invested in this subject.

The latest digitization concept [13] combines digital and operational data from industrial assets with a software, simulation and analysis platform for information on current and future operations. The result is improved production, reduced costs, and accelerated innovation. This concept encompasses the organization's digital information mix with live streaming data of how the plant is operating, understanding operational failure modes and avoiding unplanned downtime, improving performance, factory.

B. Model-based Systems Engineering and Industry 4.0

Since the beginning of the fourth industrial revolution in the 2010’s, many industrialists and researchers have worked in order to propose an architecture for data analysis. Only few proposals have been developed using model-based systems engineering (MBSE). As highlighted by Wortmann et al. [14], MBSE is a key enabler for building complex systems of systems, which can contribute for successfully engineering Industry 4.0 systems of systems. Some of these proposals are briefly described below.

Manoj Kannan et al. [15] follow a model-based requirements engineering technique in order to obtain a standardized development of information systems compliant with the manufacturing Reference Architecture Model for Industry 4.0 (RAMI 4.0) [15]. Durão et al. [17] present a preliminary integrated component data model in UML for the implementation of CPPS and Smart Product features. Petrasch and Hentschke [18] propose an Industry 4.0 process modeling language (I4PML) that extends (UML profile with stereotypes) the OMG’s BPMN (Business Process Model and Notation) standard and describes a method for the specification of Industry 4.0 applications using UML and I4PML.

These previous contributions of MBSE for Industry 4.0 partly describe a new architecture (a part of the process or a specific technology). This paper proposes to build a more complete and advanced manufacturing platform using MBSE.

III. METHODOLOGY

The SysML standard [19] defines a notation, not the way of using it. SysML therefore needs to be associated with a methodology. Accordingly, the paper adopts the phases and milestones defined by [15] in accordance with systems engineering concepts. In Fig. 1, Phase 0 refers to mission analysis and needs identification. Phase A is a feasibility study that contains system concepts and evaluates technical and programmatic aspects. Phase B establishes a preliminary design definition confirming technical solutions using trade-off studies. During Phase C, activities establish a detailed project, fabricate in preparation for integration. Purchase refers to acquisition of what is necessary for the new system. During Phase D1, activities are performed to assemble, integrate, test, and start the system. During Phase D2, preliminary tests evaluate first system performance data as preparation to assisted pre-operation tests at Phase E. The latter assesses system performing in comparison to functional requirements. Phase F refers to the closure of the project, with production of reports.

![Fig. 1. Phases of project development.](image)

IV. SYSML AND TTool

TTool [7] [6] customizes the OMG-based SysML [21] to meet the needs of real-time systems modeling. TTool supports a method. The requirement capture step reuses SysML requirement diagrams and adds modeling assumptions diagrams to explain how the model simplifies the system. Analysis is use-case driven and documented by scenarios (sequence diagrams) and flow-charts (activity diagrams). Finally, the design step defines the architecture of the system as a set of blocks and express the behaviors of the blocks as timed state machines.

The simulator of TTool animates SysML design diagrams and enables early debugging of SysML models. The model checker and the verification by abstraction module of TTool deeply explore the behavior of the SysML model, as soon as the state space of the latter is finite.

Like many SysML tools, TTool is particularly adapted to modeling the control part of real-time systems and to analyze their models in terms of temporal ordering of events. Next sections identify its limitation to address systems where the data part supersedes the control part.

V. SYSTEM MODELING

The use-case diagram in Fig. 2 applies to a broad variety of system models. Two actors, Administration and Engineer, pick information. The system takes charge of the necessary data and analysis to be presented. The engineer changes parameters; he or she performs updates and maintenance.

![Fig. 2. Use case diagram.](image)
Fig. 3 uses a block diagram to depict a context diagram and to show the system components on the production floor (industry), cloud, and elements to control and present information (see next section for further details).

A Block Diagram depicts the architecture of the system. In Fig. 4 and subsequent simulations, data gathering on the production floor and information presentation is missing since it repeatedly and automatically occurs. Inside the blocks, the messages exchanged are numbered using three letters and three numbers and they reflect the origin and purpose of the message. For example, RQM001 is a request from the Administrator to a Mobile asking for a specific data and, due to this request, the Mobile sends the message IND001 to the API in order to start an Analysis by using the message ALE002. All requests are acknowledged by using the same three letters and increasing the number of the original request.

The Administrator may select information using either his/her Mobile or a PC. In Fig. 5 the administrator selects information using the Mobile and issues RQM001. Following this initial request, Mobile sends an order to Application Programming Interface – API (IND001) and, after data gathering from Big Data and Analysis, API sends in turn an alert status message (ALE001) and information (IND002) back to Mobile to be shown in all devices, ending the cycle.

VI. GENERAL ARCHITECTURE

As a result of the MBSE prototype developed in the previous section, this section presents the general architecture of an advanced Industry 4.0 platform. This section demonstrates how data flow from the sensing of productive lines, through data analysis. It also presents results with greater value added to users (engineers, administrators, technicians).

A. Architecture of the Advanced Industry 4.0 Platform

Fig. 6 presents the global architecture of the proposed advanced manufacturing platform.

Three parts are represented: i) industry existing equipment, ii) Big Data/Analytics in cloud computing and iii) user interface. The technologies developed for the two last parts are presented below.

Big Data: Data is collected in industrial Programmable Logic Controllers (PLCs) deployed on the lines and sent and stored in the cloud with a Big Data solution. The data from the
PLCs will be constantly sent to the Big Data that will store them, so the system will contain both the current data and the history of the data already collected to be used in the analyzes.

**Analytics**: The introduction of a Big Data analysis system that allows one to feed a system of intelligence, optimization and decision support aid. Analysis results will be available in a cloud API to be consumed by front-end applications (Web and Mobile). The solution should process the data and provide results for line improvements. These improvements can be validated through POCs.

**API**: A back-end application system that will be running on a server providing storage, security, and access to user data and Analytics results. The system will also encode the daily reporting services with Analytics results and alert services for Mobile applications.

**Web**: A front-end application system that will be accessible to users via the Internet through any web browser. User control services (register, remove and edit) and control and parameterization of the Analytics system will be provided. Users of the system will also view the Dashboard that visually presents results of Analytics.

**Mobile**: Such an application can be installed on mobile devices (mobile phones and tablets). Users will be able to track Analytics real-time results, receiving alerts via push messages about issues and anomalies that are detected by Analytics.

**Dashboard**: This is a feature accessible through the Web system and can be viewed in real time by PCs, notebooks, tablets, and displayed on big screens and smart TVs.

**Virtualization**: It composes the integration between the virtual factory and the actual factory. The data collected from the real factory are used as inputs to the virtual model, making it possible to analyze the factory behavior and generate hypothetical scenarios. Virtualization allows one to validate the results anticipated by optimization and decision-making before the actual factory deployment.

The data is constantly extracted from the PLCs and sent to a local database (local DB), from there a routine installed in the local database sends the data to a Big Data system that will contain capacities to store and process large volumes of data. Big Data will do all the necessary transformations in the data to improve processing, and keep a copy of the current and historical data. This ensures isolation so that the proposed solution does not interfere with other industrial production systems.

Analytics will be constantly analyzing the data and generating results (data with higher added value). The results are sent to an API, which is a service-based system. The main functions of Analytics are: execution of artificial intelligence algorithms such as machine learning, pattern detection, prediction and optimization; calculation of indicators and graphs as production, yield, efficiency and quality; generation of alerts by detection of anomalies, failures, prediction of problems; generation of daily, weekly, monthly reports, given by the grouping of diverse information that contains relevance and context.

The API application has its own database and does more routine activities such as controlling user data, storing analytics parameters, sending push alerts to Mobile applications, sending reports, storing results, and analytics indicators. Overall, the API provides all of the key services for front-end applications: Mobile and Web.

The Web application will interface with users through web browsers (Browsers) running on PCs and Tablets. Contain login screens, user control: register, remove, edit; visualization of graphs and indicators (Dashboard screen); system control and Analytics parameters.

Users registered in the system can access the platform through various devices: PCs, notebooks, tablets, smartphones, big screens and smart TV. Access is controlled through login and password. The web system is designed to be accessible from any browser (through PCs, notebooks, tablets and smart TVs). The Dashboard is displayed within the web system as a dashboard that concentrates all the most relevant information on a screen that will have appropriate display mode for screens and smart TVs. A mobile app for smart phones will also be available. Through this application you can receive event alerts: problems, anomalies, failures, prediction problems. There will also be login screens and simplified dashboards with real-time indicators.

The growing mass of data that is worked on daily by companies brings with them immense opportunities to leverage business-related decisions and can offer competitive advantage in the corporate sector. Industrial PLC data, RFID tags, sensors, cell phones and smart meters are driving the need to deal with huge amounts of data in real-time.

### B. Big Data/Analytics Solution

Given that the large amount of data that can be collected as well as speed processing, tools often referred to as Big Data are needed to efficiently handle the data. There are several architectures for developing Big Data systems, as well as various pieces of software available to implement them. The most commonly used architectures today are Lambda and Kappa. The most common software for implementing these architectures for Big Data is Apache Hadoop and Apache Spark.

Compared with the Lambda architecture, the Kappa architecture is simpler to implement because it has only one processing layer (the Lambda architecture is divided into three layers), and is created for the sole purpose of processing data streams. In addition, its maintenance and debugging will be simpler for the same reason. The Kappa architecture was initially chosen in this project because it is a simpler architecture and meets the needs of the project. The Kappa Architecture will enable real-time analysis without loss of relevant information and history storage, which is important in the industry context. This decision led to choose a streaming processing technology and an appropriate broker.

For the Broker service, Kafka was chosen as a producer/consumer service. Kafka abstracts the flow of messages into threads, where producers (devices that send data) record the data in topics and consumers (other clusters) read the data from those topics.
For Streaming processing, Spark is used. Apache Spark is an in-memory data processing system designed to accelerate the processing of large amounts of data. Its architecture was designed for parallel clustering. It provides modules for streaming processing, module for creating machine-learning algorithms, modules for SQL queries and modules for creating algorithms based on graphs. With Spark there is the possibility of joins in tables, based on the columnar model of the Apache Cassandra and Apache HBase DBMS. The Spark still facilitates, if necessary, the change to the Lambda architecture, by having batch processing.

The Spark Streaming mode allows organizations to get information about the data at the last minute or the last hour. These data can be consumed in SQL, or NoSQL databases that feed control panels in real time, or are used in machine learning models to create indexes for doing research and gain insight into Analytics. As a NoSQL tool, Cassandra was used, which provides a scalable and resilient operational database for real-time analysis. For a better understanding of the communication between the tools described, Fig. 7 details the interaction of the Big Data and Analytics solution in a more specific way and in conjunction with the other parts of the system.

Initially the data received via representational state transfer (REST) calls and placed in Kafka topics, which distributes the data in partitions to be parallelized. Partitions are consumed in Spark, which filters and aggregates them through Spark. Spark includes map, reduces, and aggregates operations through the resilient distributed datasets – Hadoop / Spark data abstraction. Finally, the resulting dataset is saved in Cassandra. Spark's communication with Cassandra enables machine learning through the Spark Machine Learning Lib library (Spark MLlib) for statistical analysis, predictions and can be easily coupled to the Analytics system. Analytics can also take advantage of a number of other machine learning, optimization and data mining techniques for its operation, as Artificial Neural Networks (ANNs), Genetic Algorithms (GAs) and Knowledge Discovery in Databases (KDDs). Information extracted from the query in Spark is processed by Analytics and the results are sent to the API.

By examining large amounts of data to discover hidden patterns, unknown correlations, and possible predictions, Big Data Analytics is often used. The potential of Analytics is realized when the decision-making process is leveraged through its use. Increasingly, companies are looking for efficient ways to turn large and varied data volumes into powerful insights. According to the consulting firm Gartner (see Fig. 8), this segment can be divided into four distinct levels of analysis: descriptive, diagnostic, predictive and prescriptive, as shown below.

The overall architecture developed establishes a simple, feasible and low initial investment for the development of proofs of concept (POC) for the industry. Aiming at a system in which it can evolve in the future for all types of analysis presented. However, there are limitations that depend both on technical limitations in algorithms in the current literature and on the lines in which POC will be implemented. The technical limitations are mainly due to the processing time, the number of variables to be considered and the accuracy of the algorithms. The limitations of the line are mainly in the automation and quality of sensed data, in general the less automated and inaccurate the line, the more difficult it is to use techniques providing information with higher added value.

C. Results of the User’s Interfaces

Analytics results will be displayed on the controller/dashboard, mobile and web in the form of graphs and notifications (Fig. 9). The graphs selected will be tailored to the needs of the company and an in-depth study of the data that will be carried out during the analytics construction steps.

By examining large amounts of data to discover hidden patterns, unknown correlations, and possible predictions, Big Data Analytics is often used. The potential of Analytics is realized when the decision-making process is leveraged through its use. Increasingly, companies are looking for efficient ways to turn large and varied data volumes into powerful insights. According to the consulting firm Gartner (see Fig. 8), this segment can be divided into four distinct levels of analysis: descriptive, diagnostic, predictive and prescriptive, as shown below.

The overall architecture developed establishes a simple, feasible and low initial investment for the development of proofs of concept (POC) for the industry. Aiming at a system in which it can evolve in the future for all types of analysis presented. However, there are limitations that depend both on technical limitations in algorithms in the current literature and on the lines in which POC will be implemented. The technical limitations are mainly due to the processing time, the number of variables to be considered and the accuracy of the algorithms. The limitations of the line are mainly in the automation and quality of sensed data, in general the less automated and inaccurate the line, the more difficult it is to use techniques providing information with higher added value.

C. Results of the User’s Interfaces

Analytics results will be displayed on the controller/dashboard, mobile and web in the form of graphs and notifications (Fig. 9). The graphs selected will be tailored to the needs of the company and an in-depth study of the data that will be carried out during the analytics construction steps.

By examining large amounts of data to discover hidden patterns, unknown correlations, and possible predictions, Big Data Analytics is often used. The potential of Analytics is realized when the decision-making process is leveraged through its use. Increasingly, companies are looking for efficient ways to turn large and varied data volumes into powerful insights. According to the consulting firm Gartner (see Fig. 8), this segment can be divided into four distinct levels of analysis: descriptive, diagnostic, predictive and prescriptive, as shown below.

The overall architecture developed establishes a simple, feasible and low initial investment for the development of proofs of concept (POC) for the industry. Aiming at a system in which it can evolve in the future for all types of analysis presented. However, there are limitations that depend both on technical limitations in algorithms in the current literature and on the lines in which POC will be implemented. The technical limitations are mainly due to the processing time, the number of variables to be considered and the accuracy of the algorithms. The limitations of the line are mainly in the automation and quality of sensed data, in general the less automated and inaccurate the line, the more difficult it is to use techniques providing information with higher added value.

C. Results of the User’s Interfaces

Analytics results will be displayed on the controller/dashboard, mobile and web in the form of graphs and notifications (Fig. 9). The graphs selected will be tailored to the needs of the company and an in-depth study of the data that will be carried out during the analytics construction steps.

By examining large amounts of data to discover hidden patterns, unknown correlations, and possible predictions, Big Data Analytics is often used. The potential of Analytics is realized when the decision-making process is leveraged through its use. Increasingly, companies are looking for efficient ways to turn large and varied data volumes into powerful insights. According to the consulting firm Gartner (see Fig. 8), this segment can be divided into four distinct levels of analysis: descriptive, diagnostic, predictive and prescriptive, as shown below.

The overall architecture developed establishes a simple, feasible and low initial investment for the development of proofs of concept (POC) for the industry. Aiming at a system in which it can evolve in the future for all types of analysis presented. However, there are limitations that depend both on technical limitations in algorithms in the current literature and on the lines in which POC will be implemented. The technical limitations are mainly due to the processing time, the number of variables to be considered and the accuracy of the algorithms. The limitations of the line are mainly in the automation and quality of sensed data, in general the less automated and inaccurate the line, the more difficult it is to use techniques providing information with higher added value.

C. Results of the User’s Interfaces

Analytics results will be displayed on the controller/dashboard, mobile and web in the form of graphs and notifications (Fig. 9). The graphs selected will be tailored to the needs of the company and an in-depth study of the data that will be carried out during the analytics construction steps.
effectiveness (OLE) that is an aggregation of all equipment’s OEE and enables to assess the efficiency of an entire production line. Technicians and engineers will also receive push when best settings of the production line are available.

In future work CPS can be developed in a production line and directly set the parameters line with the analytics results.

VII. CONCLUSIONS AND FUTURE WORK

In the new era of advanced and digital manufacturing, the analysis of data coming from the shop-floor is a big challenge that involves a set of new technologies of the Industry 4.0. In order to contribute to this topic, this paper provides a new generic and replicable platform for advanced manufacturing in the Industry 4.0 context. The project uses a Model-Based methodology and models the system under design in SysML. The SysML models enable traceability analysis i.e., in case of a technology get updated or removed, then the tools that analyze these models will help in tracing such modifications. Thus, one of the main reasons why a model-based methodology was chosen is to achieve consistency in our work and foster safety and quality in the designed models.

The proposed platform enables to help users in their decision-making, to identify the most responsible equipment for loss of productivity and to increase the traceability of the processes. The next step toward advanced manufacturing is the integration of CPS in the platform in order to have autonomous decisions at manufacturing operation using integrated machine learning.

Ongoing work also includes the development of the mathematical algorithms for particular scenarios and the implementation of the platform in three pilot industries in metalworking, textile and, food and drinks sectors.

Acknowledgements

TTool has been developed by Dr Ludovic Apvrille, Telecom ParisTech, Sophia Antipolis, France.

REFERENCES