A technological demonstrator for cloud-based performance monitoring and assessment of industrial plants: present architecture and future developments

Riccardo Bacci di Capaci ∗ Claudio Scali ∗∗∗ Carlo Vallati ∗∗ Giuseppe Anastasi ∗∗

∗ Department of Civil and Industrial Engineering, University of Pisa, Italy
∗∗ Department of Information Engineering, University of Pisa, Italy
∗∗∗ Corresponding author: claudio.scali@unipi.it

Abstract: This paper presents the actual status and future developments of a technological demonstrator of key enabling technologies of Industry 4.0. The final scope of the project is the realization of a platform where different solutions addressing specific companies needs can be analyzed and their performance compared. Available Industry 4.0 technologies allows the realization of various architectures, even though the most appropriate solution has to be found by the company, accounting for different aspects, as economy, security, specific skills to be maintained or given in outsourcing.

The first core of the platform, already working, consists in a Control Loop Performance Monitoring (CLPM) system which operates in cloud as a single module able to supervise routine data coming from different plants. This is an attractive solution for many companies allowing to avoid costs of local installations, linked to monitoring systems, human resources and additional effort for maintenance and upgrading. Some technical details about the application on a pilot-scale plant are given to illustrate the status of the activities. The scope of the project is to add new features to the system as: CLPM extended to plants located in different sites, equipment condition monitoring and environmental data analyses. To this aim, future developments of the platform are discussed in terms of improved technologies, different protocols and architectures.

Keywords: Performance assessment; industry 4.0; cloud-based monitoring; fog/cloud computing

1. INTRODUCTION

Industry 4.0 is definitely a current hot topic since both academia and industry are rushing to get involved in order to exploit the impressive characteristics and large opportunities offered by what most players even consider a proper industrial revolution. Paradigms of Industry 4.0 were born and firstly spread, especially in Europe and Italy, mainly within the manufacturing sector. Nevertheless, strong and integrated automation naturally gives also large advantages to all the process industry. Significant returns of investment for entire industrial plants can be obtained by the implementation of technologies, algorithms and methods related to such paradigms. In Italy, also attractive fiscal benefits offered by last governments to virtuous companies investing in Industry 4.0 are giving a remarkable impulse to this digital transformation.

Among the manifold theoretical definitions, a practical approach is to consider at Industry 4.0 level any physical and digital solution which adopts at least one of the Key Enabling Technologies (KET): Advanced Manufacturing Solutions, Integrated Simulation, Industrial Internet of Things (IIoT), Cloud, Cyber-security, Big Data Analytics (BDA). Cloud Computing, defined by Mell and Grance (2011), allows the realization of attractive process-related applications for many scopes such as: data historians; analysis tools; alarm, asset and performance management; training simulators; remote diagnostics (Latha and Jayaprakash, 2017). In particular, cloud computing is of primary importance for smart industry with objectives of process monitoring, control and optimization.

It is to be noted that various cloud computing platforms for IIoT and BDA are commercially available. Key players are really variegated: cloud services providers (Amazon Web Services, Microsoft Azure, Google, Intel, IBM), enterprise solution vendors (as Oracle and PTC), networking companies (like AT&T, Verizon, Cisco) and industrial engineering companies (e.g., Siemens, ABB, AspenTech, Metso, Rockwell Automation, Honeywell, Bosch and General Electric), among others. Some platforms require a proprietary licenses; few of them are accessible as open source projects (Kabugo et al., 2020). However, the optimal cloud-based solution must be chosen by the company with reference to its specificity. Therefore, there is space to analyze general architectures and compare different solutions. This is also the conclusion of a very recent survey on this subject, including different definitions of cloud, comparison of on-premises vs. off-premises solutions, pro and cons of various levels of services offered by vendors of different platforms (De Caigny, 2019).

Nevertheless, examples of implementation of data analytics in the context of cloud computing and Industry 4.0 with purposes of process monitoring, control and optimization are still scarce in the scientific literature. Among the few applications, an interesting industrial cloud-based architecture applied to an elec-
tric induction motor was developed by da Silva et al. (2016). Another cloud-based condition monitoring system for machinery with application to power plants was recently presented by Elazab et al. (2017). Moreover, Kabugo et al. (2020) presented an IIoT software platform, based on cloud computing, big data analytics and machine learning, able to guarantee data acquisition, employ statistical methods for data pre-processing, develop soft sensor, and perform real-time process monitoring and offline data analytics. In particular, data-driven soft sensors were derived to predict syngas heating value and hot flue gas temperature in a waste-to-energy (WTE) plant.

The contribution of this work is thus to present a technological demonstrator for global monitoring and assessment of process plants taking advantage of various Industry 4.0 technologies. The paper has the following outline: Section 2 illustrates the general structure of the proposed technological demonstrator, a cloud platform for performance monitoring and assessment. Section 3 presents the actual status, with some details of the working CLPM module and of the pilot plant used as first demonstrator core of the whole system. In Section 4, future developments in terms of alternative cloud architectures and additional technologies to be soon implemented are discussed. Finally, concluding remarks are reported in Section 5.

2. THE TECHNOLOGICAL DEMONSTRATOR

The first core of the cloud-based platform is already working and concerns a Control Loop Performance Monitoring (CLPM) module. The actual system operates as single unity and allows remote supervision of PID control loops of a pilot plant.

This system will be enhanced in the close future with three novel analytics tools: i) extended CLPM: the first cloud-based solution for CLPM will be applied to different plants located in distant areas to test the system scalability; ii) equipment condition monitoring: other physical machines, e.g. electrical engines, pumps, compressors, to perform energy consumption and vibration analysis will be analyzed with the aim of condition monitoring and predictive maintenance; iii) environmental analysis: variables, such as temperature, gas and liquid concentrations, will be supervised to perform specific data analytics. In parallel, some hardware and software upgrading will be studied: mainly, the implementation of fog computing, an improvement of IoT integration, and the adoption of data mining algorithms for predictive maintenance. The result will be a comprehensive cloud-based system, a Platform as a Service (PaaS), for global monitoring of industrial process plants.

By providing low-cost and abundant computing and storage resources, cloud technologies enable Advanced Manufacturing solutions which exploit the analysis of large amount of historical data through complex techniques, known as Big Data Analytics (BDA), whose implementation is actually unfeasible with the limited resources available in traditional systems. Another major advantage of cloud computing is that centralized systems can be applied in localized remote servers. Data from different industrial plants and separated production sites can be transferred and analyzed in cloud by a single monitoring system. This can be a particular appealing solution when control engineers have to monitor a large set of similar plants or units, and, consequently, have to tackle common issues and faults, or when they have to supervise and/or control geographically dispersed assets, as for water or energy supply plants (McGrav, 2018). In addition, cloud-based systems have natural attractive features, being easy accessible, replicable, distributable, and adaptable. For example, the use of an unified logic highly reduces risks of incorrect replications on single systems and involves minimal employment of routine plant staff, since the whole task can be carried out by dedicated and qualified company units or can be outsourced as a service to specialists of third-companies.

Nevertheless, our long experience on performance monitoring, and CLPM in particular, makes us believe that this process may be long and tortuous. Practical difficulties may arise in transferring skills acquired over years by operators on singles plants, since specific competences of processes and operations cannot be fully generalized and may not be exported quickly. In addition, since process data are confidential, industrial companies may be not trustful of moving data from local computer systems to external clouds, as additional issues of cybersecurity and service reliability arise. Therefore, industrial companies may prefer to own and be in charge of their own clouds.

3. THE ACTUAL SYSTEM

In this section, the actual cloud-based architecture, with details of the solutions adopted for data acquisition from the field and transmission to the cloud are presented.

3.1 The CLPM module

Since at least two decades, major industrial engineering companies have offered their own “traditional” solution for CLPM, that is, an on-premises software tool, which has to be installed within the local computer systems of the different industrial plants under supervision. The most recent comparison of CLPM packages has been carried out by Bacci di Capaci and Scali (2018), within a review of literature on valve stiction. In the last years, several industrial and control engineering companies have moved their on-premises software in the cloud. Among others, Siemens promotes a specific module for CLPM within MindSphere, its PaaS for Industrial IoT.

Within our cloud-computing platform, the first established application for process monitoring is devoted to CLPM. This analytics tool exploits the PCU system, a long-standing software developed in the Chemical Process Control Laboratory of University of Pisa. Historically, the first full release of the system with large industrial implementations was discussed in Scali and Farnesi (2010). Updated versions with further large-scale applications were then reported (Bacci di Capaci et al., 2013; Bacci di Capaci and Scali, 2014).

Our CLPM system, under continuous updating, can diagnose sources of malfunction of traditional single-input single-output (SISO) control loops with PID controllers, by suggesting actions to be taken. External disturbances, poor controller tuning, faults of valves and sensors, and interactions with other loops are the main detectable malfunctions. These issues cause oscillations in the process variables and therefore their distinction is of primary importance in order to carry out the most suitable correction. The availability of different measurements from the field implies the activation of various analysis paths. For example, when valve position is registered, as Fieldbus communication and smart positions are used, an enhanced actuator diagnosis can be performed. Details about core techniques and algorithms installed into the various analytics modules, as well as specific examples of application are reported in Scali and Farnesi (2010) and Bacci di Capaci et al. (2013).
Within traditional systems for CLPM various modules interact with each other and are all physically implemented into the local computer control system. Interface, assistant, and application are typically the three main functional categories which the various modules belong to (Lee et al., 2010). Also our traditional CLPM systems have these features. The architecture of the first version of PCU system, currently monitoring several refinery plants, can be found in Scali and Farnesi (2010). For the later on-premises version (PCU+), a system specifically devoted to smart industries and running in several power plants, the architecture of Figure 1 is employed. A Scheduler module leads data acquisition and processing operations; various OPC servers and corresponding OPC clients running acquisition applications are employed to collect real-time data; a Viewer application is used to show analysis results with plots and verdicts. An interface of the database (DB) is edited to perform loop configuration. The PCU analytics tool, developed in MATLAB, works as an executable program and exchanges input/output data once activated by the Scheduler.

### 3.2 The Actual Architecture

The CLPM module of our cloud-based platform exploits a standard centralized architecture (see Figure 2). No module is indeed on-premises, but all elements are installed within a remote Linux cloud server. Once transmitted to the cloud, data are then written and stored into the cloud database, which acts as a star center of the whole network. An updated analytics tool (PCU-Cloud) is the core of the whole CLPM module and runs once activated by a light scheduler. A Java web interface queries the database and allows result visualization, as well as many aspects of loops setting and all phases of analysis management. Note that, when peculiarities of industrial sites may need to be completed within local computer systems. The web application is accessible at the following link: https://pcu.rjsoft.it/pcu-cloud/. Any interested guest user can access with reader level and manage the application by entering username and password. The main functionalities of the web app are illustrated in Bacci di Capaci and Scali (2020).

The CLPM module employs standard features of cloud technology: JSON data format and MQTT protocol. JSON (JavaScript Object Notation) is a data interchange format, now very spread in IoT; it is ISO-standard, based on TCP/IP, and holds intrinsic cybersecurity features. The CLPM module uses JSON data format and MQTT protocol. JSON (JavaScript Object Notation) is the most common and interoperable messaging system for IoT; it is ISO-standard, based on TCP/IP, and holds intrinsic cybersecurity features. According to MQTT implementation or definition. Five PID control loops are now programmed into the PLC; four pressure loops with pneumatic actuator, one flow rate loop with electric valve. PLC operates each loop with a period of 1 s, collects and stores data in its local DB, and then transfers data to the cloud server. The various

### 3.3 The Pilot Plant

IdroLab is a pilot plant located at the Consorzio Polo Tecnologico Magona of Cecina, Livorno (Italy). The plant has been recently upgraded to become the first demonstrator facility of Industry 4.0 technologies in the framework of a project developed by CLUI AS, the Italian Association of Automation & Instrumentation End Users. A novel PLC (Siemens, Simatic S7-1500) is installed and configured with Simatic Step7 TIA Portal program to control operations. The plant is comprised by two hydraulic circuits equipped with a centrifugal pump under inverter control (see Figure 3). Process operation and variables measurement (pressure, flow rate and level) are possible by a set of latest generation actuators and sensors. Profinbus protocol is used for Fieldbus communication from the smart devices to PLC; while MQTT protocol is employed for communication from PLC to cloud server.

The plant is being used as first field of application to test our cloud-based process monitoring platform. The CLPM module has been extensively tested and now is perfectly working. Other hardware and software functionalities are currently under implementation or definition. Five PID control loops are now programmed into the PLC: four pressure loops with pneumatic actuator, one flow rate loop with electric valve. PLC operates each loop with a period of 1 s, collects and stores data in its local DB, and then transfers data to the cloud server. The various

---

**Fig. 1.** An example of on-premises architecture for CLPM: the PCU+ system (Bacci di Capaci et al., 2013).

**Fig. 2.** The cloud-based solution with only the CLPM module.

**Fig. 3.** IdroLab: the testbed for the cloud-based platform.
data are written within a defined string in JSON format with key and value notation. The following data are transmitted: name, time stamp, loop measurements – set-point, process variable, PID output, as mandatory data, manipulated variable (that is, valve position), valve positioner drive signal, electropneumatic converter output, as optional, and PID controller parameters – proportional gain, integral constant, derivative constant, filter constant, controller mode, high and low limit on variables.

Main causes of fault in PID control loops can be reproduced by using some physical modular items as described in Scali et al. (2011). In addition, malfunctions can be also introduced by the use of dedicated software blocks. For example, valve stiction is reproduced by activating a data-driven model within a customized function block of the PLC. External software disturbances are introduced within the inverter and two motored valves. A sinusoidal disturbance can be imposed to the desired inverter velocity, by setting amplitude and frequency of oscillation around the set-point value. Moreover, the input signal to motored valves can be altered with a step-wise wave.

4. POSSIBLE FUTURE ENHANCEMENTS

We are currently updating the proposed cloud-based platform and the IdroLab testbed in order to introduce other monitoring applications, improve current functionalities and extend the interface. Specifically, we aim at enhancing our work in the following directions: i) improve the architecture of the platform in order to ease the large scale adaptation of the solution for CLPM and test the scalability of the system by integrating additional plants located in different areas; ii) integrate other physical equipment, such as electrical engines, pumps, compressors, to perform energy consumption and vibration analysis; iii) integrate sensors to monitor environmental variables, such as temperature, gas and liquid concentrations. In the following, we will provide an overview of the possible future enhancements for the platform to implement in practice the above mentioned functionalities.

4.1 Extending the Cloud Architecture

The first architecture for CLPM module depicted in Figure 2 is a standard cloud-based solution, tested over IdroLab pilot plant. However, it has to be noted that all the messages directly published by the various field elements can be naturally managed by the MQTT broker, regardless the element type, even when installed in geographically far industrial sites. Therefore, the first architecture can be easily expanded as shown in Figure 4. Different plants can be thus monitored by an unique fully centralized cloud system. DCS, PLC, or single smart devices, as sensors and actuators, of different plants, production sites and even industrial companies can directly communicate with the same MQTT broker.

In this solution, one can ensure high levels of cybersecurity: the server provider guarantees intrinsic safety of the cloud platform; a three-level security system is used to protect communication channel: a secret alias for the broker, and unique username and password to identify various clients. In addition, by turning to MQTTs (where S stands for Secured), protocol security can be enhanced. This means to adopt a certificate on server side during TLS handshake, which avoids “man-in-the-middle” issues. Note that additional safety requisites of local networks, in terms of firewall, protected ports, etc. are still under the responsibility of the various industrial companies.

This extended platform is programmed to be scalable and industry oriented. The software side does not cause additional issues, as core codes are general and can be easily exported to different scenarios. When one has to be monitor numerous sites or large plants with thousands of elements, the number of messages published into the cloud may become high, therefore physical architecture limitations need to be overcome, by augmenting transmission band, CPU, RAM and so on. Otherwise, during configuration of local client systems, one may reduce transmission frequency and then data size to limit data traffic and save cloud space. However, it is important to note that such fully centralized cloud architecture cannot be the final solution for a global monitoring platform.

4.2 Implementing Fog Computing

Despite the many advantages, the Cloud technology is indeed characterized by limitations and disadvantages. A fully centralized architecture, as the one of Figure 4, confines data and functionalities into data centers which can be very distant from the physical locations where information are produced. This centralized architecture requires the data to be transmitted through an Internet connection, which can represent a significant drawback for critical applications (Abdelwahab et al., 2014), like some industrial scenarios with specific requirements in terms of confidentiality and reliability. This continuous offload of the data can represent a confidentiality concern and an unacceptable point of failure for some applications, as an interruption of the Internet connection results in a break of the flow of data and its analysis. In addition, the Internet introduces a significant delay, which might not be tolerable for critical applications, e.g. the ones that require a timed data processing for monitoring.

In order to overcome such limitations, Fog computing can be adopted. This technology introduces an intermediate computing and storage layer between the cyber physical systems and the Cloud infrastructure (Bonomi et al., 2012). The Fog layer is deployed directly into the physical locations where the data is produced and aims at supporting the execution of critical applications. This additional layer is implemented by means of high performance embedded systems, the Fog nodes, which are in direct communication with PLC, sensors or other devices.
A future evolution of our global performance monitoring platform is to adopt the Fog computing paradigm. The centralized architecture can be extended by installing a Fog node in proximity of the various analysis modules (or co-located with them) to support the execution of a part of the functions of the system. The runtime environment provided by each Fog node could allow to migrate some of the functionalities from the Cloud to the Fog in a seamless manner without requiring modifications to the current implementation. The possibility to execute some of the functions directly in proximity of the plant could enable functions that require real-time analysis, since the direct communication with the physical systems allows to collect and analyze data within predictable time limits. Big data analysis on historic series can still be implemented on the Cloud as data can be offloaded after a first analysis, as shown in Figure 5.

The possibility to run some of the functions directly on site could help in improving security and confidentiality. Sensitive data could be analyzed locally without the need to move them to the Cloud, while data anonymization and filtering could be implemented locally before sending the data to the Cloud in order to enforce confidentiality. These functionalities could be beneficial to foster the adoption of the system on a large scale: the improved security and confidentiality is expected to convince industrial companies to adopt the system, whether they were not favorable to standard cloud-based solutions.

4.3 Improving IoT integration

The current implementation of our platform exploits the MQTT protocol to collect the data from the devices and forward them to the cloud. The core of the system is the MQTT broker that is responsible for receiving all the communication from the sensors and dispatching them to the Cloud application. Although this solution is widely adopted today in many commercial systems, it has some limitations. The main drawback is the broker, that has to handle every message generated by the devices. A system that integrates different plants or even various companies is expected to handle a large amount of messages. In this scenario, MQTT can represent a performance bottleneck since all the messages are transmitted to the broker. The result is that the maximum collection frequency that can be adopted on a site could be limited by the available bandwidth provided by the Internet connection used to transmit the data.

The Fog computing paradigm can help in mitigating this issue: by moving some of the functionalities directly on the plant, (a part of) the data could be analyzed locally without offloading it through the Internet. The centralized MQTT approach, however, is not well suited for a multi-layered architecture, where data can be dispatched to both the cloud and local Fog nodes. In order to implement a multi-layer architecture, a different solution based on the Web of Things paradigm has been proposed recently (Guinard and Trifa, 2009). The Web of Things integrates physical devices into Cloud platforms through the same client/server approach originally defined for the World Wide Web. Rather than sending the measurements continuously to the cloud, each device is programmed as a server to expose an interface which can be invoked via network connection to retrieve the data by any other element implementing the functionalities of a client. In this way, an application running in the Cloud or in a Fog node can retrieve the data in the same manner a Web browser retrieves a web page from a Web server (see Figure 6). This simple model is implemented by the Constrained Application Protocol (CoAP), which has been recently standardized as the interface for the communication between applications and devices, adopting the Web of Things approach.

With respect to MQTT, CoAP can ensure a better integration with a Fog/Cloud architecture. Eventually, this approach can help to mitigate the bottleneck represented by transmission of the data over the Internet. To this aim, the Web of Things approach will be adopted in the future evolution of our platform, in order to improve its scalability and sustain both larger deployments and higher data sampling rate.

4.4 Adopting Data Mining for Predictive Maintenance

The continuous collection of data from different plants at higher rates will allow one to monitor the various status in real-time, highlighting malfunctions and issues immediately. In addition, the system will allow the creation of a large database with fine-grained time-series measurements, which can be analyzed offline to investigate the operating conditions of each part of plants. Such data analysis could help to look for hints of wear and upcoming faults, predict possible malfunctions in advance, as well as point out unconventional behaviors (Carnero, 2006).

This predictive maintenance analysis (Gouriveau et al., 2016) for monitored devices represents a significant challenge and it is of primary importance for the optimization of maintenance procedures. The ability to timely and efficiently replace malfunctioning parts can have an evident impact to reduce service outage, improve the process quality, and implement cost-effective policies. For this reason, the data collected from a plant at high rate will be stored in the cloud to create an historic repository of data, which will be used to implement novel predictive maintenance functionalities.

Specifically, data mining algorithms (Shang and You, 2019) will be developed to perform both supervised and unsupervised analysis of the available streaming data, so to timely spot out anomalies. A two step approach will be adopted in the definition of such algorithms. In the first place, a sample of real data will be collected from the IdroLab testbed in order to...
design a first definition of the algorithms and tune them. Later, the algorithms will be tested by introducing simulated faults in the system and then on data acquired from industrial plants.

To conclude, the proposed enhanced fog/cloud-based platform is shown in Figure 7. The adoption of Web of Things leads to a change of perspective: from a broker-centric solution with MQTT protocol to a sensor-centric approach with CoAP standard. The following functional applications will be performed within the Fog layer: data pre-processing (e.g., filtering, scaling, outliers removal), alarms management, critical and confidential analysis, and also traditional analytics for CLPM with PCU tool. Moreover, specific algorithms for Big Data Analytics (BDA) will be adopted into the Cloud layer within three main tools: extended CLPM on multivariate scale, condition monitoring of machineries, and environmental analyses. Finally, note that within the proposed platform data exchanges are totally bidirectional, as information and commands can be also dispatched from cloud and fog layers to the field level.

5. CONCLUSIONS

The platform here discussed is an example of remote system for global performance monitoring and assessment of process plants. Implementation details have been given, in terms of global architecture and selected protocols. Starting from a standard cloud-based solution devoted to PID control loops performance monitoring and tested over a pilot plant facility, the following enhancements are outlined: extending the fully centralized cloud solution to multiple plants of distant productive sites and of different industrial companies; implementing a fog/cloud architecture to increase mainly flexibility and reliability of the system; adopting Web of Things approach with CoAP protocol to further increase robustness and scalability; employing data mining algorithms on big data with the aim of predictive maintenance of different devices, from control loops, as sensors and actuators, to other plant machineries, as electric motors, pumps and compressors.

REFERENCES


