

Cloud-Fog Automation: The New Paradigm Toward Autonomous Industrial Cyber-Physical Systems

Jiong Jin¹, Member, IEEE, Zhibo Pang², Senior Member, IEEE, Jonathan Kua³, Member, IEEE, Quanyan Zhu⁴, Senior Member, IEEE, Karl H. Johansson⁵, Fellow, IEEE, Nikolaj Marchenko, Member, IEEE, and Dave Cavalcanti⁶, Senior Member, IEEE

Abstract—Autonomous Industrial Cyber-Physical Systems (ICPS) represent a future vision where industrial systems achieve full autonomy, integrating physical processes seamlessly with communication, computing and control technologies while holistically embedding intelligence. Cloud-Fog Automation is a new digitalized industrial automation reference architecture that has been recently proposed. This architecture is a fundamental paradigm shift from the traditional International Society of Automation (ISA)-95 model to accelerate the convergence and synergy of communication, computing, and control towards a fully autonomous ICPS. With the deployment of new wireless technologies to enable almost-deterministic ultra-reliable low-latency communications, a joint design of optimal control and computing has become increasingly important in modern ICPS. It is also imperative that system-wide cyber-physical security are critically enforced. Despite recent advancements in the field, there are still significant research gaps and open technical challenges. Therefore, a deliberate rethink in co-designing and synergizing communications, computing, and control (which we term “3C co-design”) is required. In this paper, we position Cloud-Fog Automation with 3C co-design as the new paradigm to realize the vision of autonomous ICPS. We articulate the state-of-the-art and

future directions in the field, and specifically discuss how goal-oriented communication, virtualization-empowered computing, and Quality of Service (QoS)-aware control can drive Cloud-Fog Automation towards a fully autonomous ICPS, while accounting for system-wide cyber-physical security.

Index Terms—Cloud-fog automation, industrial cyber-physical systems.

I. INTRODUCTION

RAPID technological advancements in Industrial Internet of Things (IIoT), artificial intelligence (AI), machine learning (ML), and cloud/edge computing paradigms have driven large-scale innovations in industrial automation systems in Industry 4.0 – spanning across many verticals, such as advanced manufacturing, intelligent transportation, smart logistics, smart grids, robotics, agriculture, healthcare, energy, urban planning, and so forth [1]. Modern industrial automation systems play a major role in Industrial Cyber-Physical Systems (ICPS) where physical processes aimed to be seamlessly integrated with computational and communication technologies. For instance, ICPS transforms conventional manufacturing systems into smart environments by embedding intelligence into physical processes, where machines, robots, and human operators collaborate seamlessly. These systems are characterized by their ability to monitor, analyze, and respond to changes in the environment, which ensures higher levels of operational efficiency, flexibility, and safety. However, most of these ICPS capabilities and processes are not yet fully autonomous. Therefore, there is a growing traction in academia and industry to pursue a futuristic vision of “Autonomous ICPS” where the entire ICPS holistically functions, operates and executes industrial processes with full autonomy, while simultaneously embedding intelligence throughout the system.

Autonomous ICPS will create a new technological paradigm that forms the backbone of future industrial automation systems towards Industry 5.0 [2]. The fundamental architecture of most industrial automation systems today stems from the International Society of Automation (ISA)-95 model, which has largely remained unchanged for the past two decades [3]. ISA-95 is a pyramid architecture, where all communication, computing, and control elements of ICPS (such as those in smart manufacturing enterprises) are both logically and physically partitioned into five separate levels, i.e., Levels 0 to 4. Each of the levels correspond to physical machines, low level machine control, middle level supervisory control,

Received 30 September 2024; revised 30 January 2025; accepted 28 February 2025. Date of publication 28 May 2025; date of current version 9 September 2025. The work of Jiong Jin was supported in part by Australian Research Council Linkage Project under Grant LP190100594. The work of Zhibo Pang was supported in part by Swedish Foundation for Strategic Research (SSF) under Project APR20-0023 and in part by Swedish Innovation Agency (Vinnova) through the SweWIN Center under Grant 2023-00572. The work of Quanyan Zhu was supported in part by the National Science Foundation (NSF) under Grant ECCS-1847056. The work of Karl H. Johansson was supported in part by Swedish Research Council Distinguished Professor Grant 2017-01078, in part by the Knut and Alice Wallenberg Foundation Wallenberg Scholar Grant, and in part by Swedish Strategic Research Foundation FUSS SUCCESS Grant. (Corresponding authors: Jiong Jin; Zhibo Pang.)

Jiong Jin is with the School of Engineering, Swinburne University of Technology, Hawthorn, VIC 3122, Australia (e-mail: jiongin@swin.edu.au).

Zhibo Pang is with the Department of Intelligent Systems, KTH Royal Institute of Technology, 10044 Stockholm, Sweden, and also with the Department of Automation Technology, ABB Corporate Research Sweden, 72226 Västerås, Sweden (e-mail: zhibo@kth.se).

Jonathan Kua is with the School of Information Technology, Deakin University, Geelong, VIC 3220, Australia (e-mail: jonathan.kua@deakin.edu.au).

Quanyan Zhu is with the Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University, Brooklyn, NY 11201 USA (e-mail: quanyan.zhu@nyu.edu).

Karl H. Johansson is with the Division of Decision and Control Systems, School of Electrical Engineering and Computer Science, and Digital Futures, KTH Royal Institute of Technology, 10044 Stockholm, Sweden (e-mail: kallej@kth.se).

Nikolaj Marchenko is with Corporate Research, Robert Bosch GmbH, 71272 Renningen, Germany (e-mail: nikolaj.marchenko@bosch.com).

Dave Cavalcanti is with Intel Corporation, Santa Clara, CA 95054 USA (e-mail: dave.cavalcanti@intel.com).

Digital Object Identifier 10.1109/JSAC.2025.3574587

manufacturing execution systems (MES) and enterprise resource planning (ERP) systems.

With the rapid penetration of digital technologies in ICPS, there is a concerted industrial effort towards building a more network-centric architecture, where elements in any of the five ISA-95 levels can be directly connected with elements in another other levels according to system requirements. Therefore, ISA-95 has become a logical/virtual partitioning architecture for semantic definition of functionalities instead of physical partitioning/topological layout of the industrial elements. These changes have recently led to several Cloud-based Automation architectures, such as the NAMUR Open Architecture (NOA), Open Platform Communications Unified Architecture (OPC UA), and Cloud Robotics.

Cloud-Fog Automation took this development further by proposing a novel industrial system reference architecture which significantly disrupts the conventional ISA-95 layered model by fully leveraging cloud/fog computing and AI/ML technologies to achieve the convergence of communication, computing, and control in ICPS [4]. It is a network-centric architecture where cloud and fog technologies are used collaboratively to provide determinism in connectivity, networked computing, and intelligent networked control in ICPS. Cloud-Fog Automation is a fundamental paradigm shift where industrial applications can be realized independent of vendor-specific hardware and software, and increasingly being provided as a service with varying degrees of technical and infrastructure requirements for flexible deployments (e.g., Control-as-a-Service, Robots-as-a-Service, Manufacturing-as-a-Service), which translates into significant capital expenditure (CapEX) and operational expenditure (OpEX) benefits.

An important characteristic of Cloud-Fog Automation is the deployment of next-generation wireless technologies to provide greater flexibility, redundancy, and mobility. It is imperative that these wireless technologies provide ultra-reliable low-latency communications (uRLLC) for time-critical computing and control applications, without compromising functional safety and security of ICPS. New wireless technologies such as 5G New Radio (NR), Wi-Fi 6/7, wireless time-sensitive networks (TSN) are required to guarantee the availability of the overall system, and satisfy the stringent latency and reliability requirements of Level-1 and Level-2 control. For example, when the network or cloud/fog computing infrastructure is temporarily unavailable or underperforming, the orchestration system will automatically move the time-critical tasks to closer to the data producer and consumer to mitigate the impacts.

Compared to traditional Level-1 and Level-2 (in the ISA-95 architecture) controllers, Cloud-Fog Automation offers significantly-enhanced computing power at lower costs, as driven by advancements in fog and edge computing technologies [4]. The proximity to the field and “physically distributed” architecture are crucial for reducing latency, improving reliability, and addressing security and safety challenges. The “logically centralized” and virtualized nature of Cloud-Fog Automation simplifies post-commissioning tasks such as upgrades, reconfiguration, dynamic resource allocation, load balancing, and modular task management, while

streamlining the digital distribution and delivery of products and services. Its compatibility with emerging AI technologies makes Cloud-Fog Automation an ideal platform for supporting next-generation ICPS, which are beyond the capabilities of conventional automation systems. When integrated with wireless networks, Cloud-Fog Automation further reduces installation and cabling costs, while enabling greater mobility, flexibility, and finer-grained deployments of automation systems.

Although Cloud-Fog Automation is a promising architectural proposal, there are still significant research gaps in realizing the vision of a fully autonomous ICPS. Specifically, the tight coupling across communication, computing and control domains is pivotal. Therefore, a deliberate rethink in co-designing and synergizing communication, computing, and control (which we term “3C co-design” in this paper) is required to position Cloud-Fog Automation as a new paradigm towards a fully and all-encompassing autonomous ICPS. There are currently significant gaps in understanding the intricate interplay between communication, computing, and control in this context. To this end, we have specifically identified the three targeted research domains in 3C co-design. These research domains are deemed to be the key driving forces for realizing autonomous ICPS with Cloud-Fog Automation as the transformative paradigm: (i) goal-oriented communication; (ii) virtualization-empowered computing; and (iii) Quality of Service (QoS)-aware control.

In this paper, we start by emphasizing the pressing need for 3C co-design, and clearly articulate the key constituents that are required to implement the vision of autonomous ICPS. We then provide a scoping for the state-of-the-art and latest advances in goal-oriented communication, virtualization-empowered computing and QoS-aware control; and endeavor to promote research and innovations to bridge the gap between theory and real-world applications, along with their practical implications. This paper serves as a roadmap and vision to steer research and development efforts in driving Cloud-Fog Automation as the new paradigm towards autonomous ICPS, bringing 3C co-design to the forefront.

The rest of the paper is organized as follows. Section II presents a Cloud-Fog Automation prototype with four representative use cases, outlines the design requirements of autonomous ICPS and discusses the pressing need for 3C co-design in realizing this vision. Section III introduces goal-oriented communication and its performance metrics, review the state-of-the-art, and outlines important research questions in rendering industrial communication systems goal-oriented. Section IV presents the evolution of virtualization technologies, stemming from cloud computing to industrial use cases, and specifically survey virtualization techniques in ICPS across application, communication and hardware layers, and provide future research avenues at each layer. Section V identifies the challenges of QoS-aware control in distributed control automation architectures, reviews key enabling technologies and outlines research directions. Section VI presents the critical aspect of cyber-physical security that cuts across the 3C domains, including potential security risks, attacks and defense strategies leveraging AI and large language

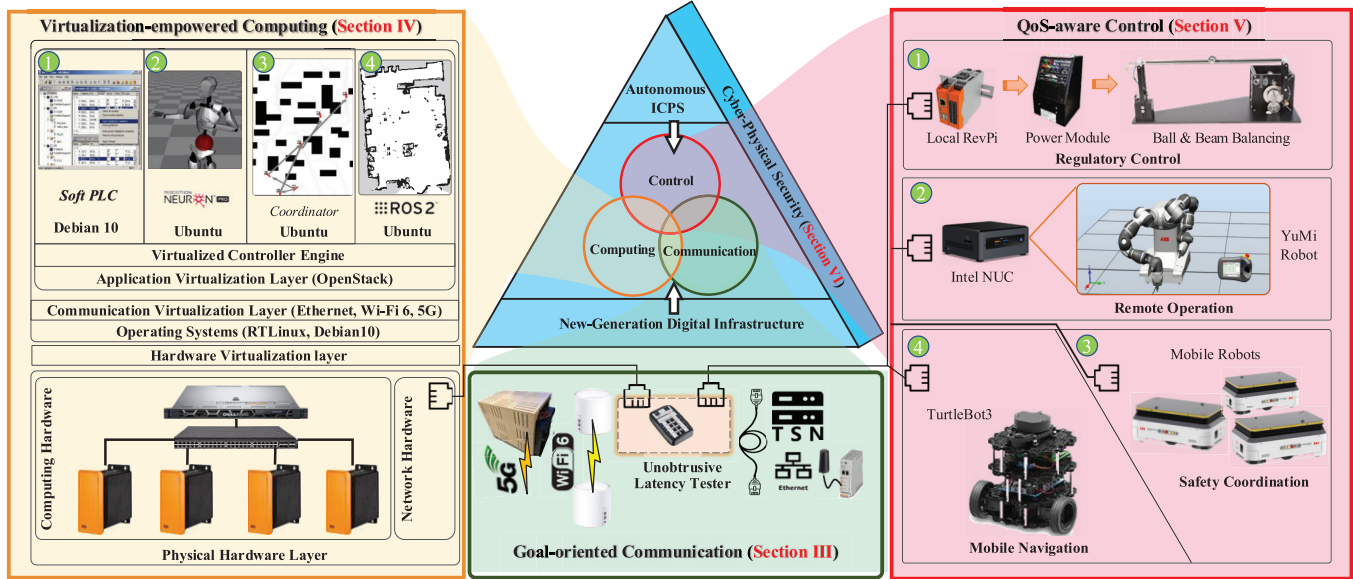


Fig. 1. Architectural overview of our Cloud-Fog Automation prototype with four representative use cases (adapted from [5]): (1) regulatory control; (2) remote operation; (3) safety coordination; and (4) mobile navigation. The central pyramid figure illustrates the crucial role of 3C co-design in closing the gap between the current/emerging new-generation digital infrastructure and the vision of autonomous ICPS. Cyber-physical security cuts across the 3C domains.

models (LLMs) techniques. Finally, Section VII concludes the paper.

II. CO-DESIGN OF COMMUNICATION, COMPUTING AND CONTROL IN CLOUD-FOG AUTOMATION

In this section, we introduce a prototype featuring four distinct industrial use cases, and outline the design requirements of autonomous ICPS. We then review technologies in current/emerging new-generation digital infrastructure, and articulate how co-designing communication, computing, and control in Cloud-Fog Automation can pragmatically close the gap to fully realize the vision of autonomous ICPS.

A. Representative Use Cases through 3C Co-Design in Cloud-Fog Automation Prototype

To illustrate the significance of 3C co-design and transformative potential in ICPS, we designed and implemented a real-world Cloud-Fog Automation prototype with four representative use cases in the industrial automation system context: (1) regulatory control; (2) remote operation; (3) safety coordination; and (4) mobile navigation. These four use cases are crucial features of modern ICPS with varying reliability and latency requirements. More specially, these four use cases present critical industrial control applications, that are enabled by the underlying communication networks and computing infrastructure. Fig. 1 presents an architectural overview of these four use cases and their respective capabilities supported by communication, computing and control technologies within the central ‘3C co-design pyramid’ figure. We investigate the impacts of current communication and computing technologies on control applications’ performance, and subsequently derive

key 3C co-design insights/design requirements for future autonomous ICPS. We refer the readers to [5] for further implementation details and result analysis.

In the first use case, the *regulatory control* process ensures compliance with industrial standards by overseeing and managing the kinematic and dynamic behaviors of systems. Cloud-Fog Automation facilitates efficient monitoring and control of time-critical industrial operations, such as continuous production processes. It enables real-time system supervision and immediate feedback to maintain adherence to regulatory standards by co-designing and integrating virtualization technologies with communication frameworks. However, challenges such as unpredictable network delays, sudden interruptions, and unstable outputs from virtualized computing can lead to production line disruptions, compromised product quality, or safety risks.

The second use case focuses on *remote operation*, enabling remote control of industrial equipment to protect workers from physically harsh and unsafe environments. Cloud-Fog Automation supports this by facilitating low-latency remote operation and real-time monitoring of on-site control states and signals. It achieves this through the deployment of soft controllers and data storage within the guest OS layer of the virtualized computing infrastructure. Nonetheless, unreliable communication and uncertainties arising from virtualization can lead to mismatches between operator commands and the actions executed by remote devices.

The third use case focuses on *safety coordination*, which ensures compliance with safety policies in industrial environments by managing the movements of multiple robots. This requires a centralized coordinator, where robots operate based on dynamic modeling and are synchronized through specific

ordering policies. Cloud-Fog Automation facilitates real-time communication among robots, enabling them to share information about the real-time positions/status of nearby objects, thus ensuring adherence to safety protocols. However, unstable communication channels may cause delays in real-time status updates or loss of control commands, potentially leading to collisions and hazardous situations.

In the fourth use case of *mobile navigation*, we evaluate the accuracy and efficiency of autonomous guided vehicles (AGVs) in navigating large-scale maps. High-speed data transmission and low-latency interactions between AGVs and edge controllers is ensured through co-designing the AGVs' control requirements with communication networks with virtualization software. Cloud-Fog Automation enables the rapid and efficient processing of large data volumes, a critical requirement for robotic navigation tasks. However, the performance of robotic control in edge computing environments depends on the specific edge computing deployment strategies [6].

In summary, the four representative use cases presented above highlight the critical importance of 3C co-design, particularly the co-design of *communication* networks and *computing* infrastructure with *control* requirements to support critical industrial applications. The presented prototype demonstrates that Cloud-Fog Automation has both the capability and flexibility to meet the stringent and multi-dimensional requirements of diverse industrial applications.

B. Design Requirements of Autonomous Industrial Cyber-Physical Systems

Although the aforementioned prototype and use cases demonstrate the significant potential of Cloud-Fog Automation, they currently lack full autonomy and require coordinated efforts to implement sophisticated autonomous industrial systems, and ultimately realize the vision of *autonomous ICPS*. The flexible and open nature of the Cloud-Fog Automation architecture offers substantial opportunities not only to enhance the autonomy of ICPS but also to optimize system performance, minimize resource consumption, and ensure resilient operations across diverse application domains. The integration of AI/ML and big data technologies within ICPS, alongside increasingly sophisticated sensing and communication capabilities, necessitates more than just timely data delivery and reliable communication. Here we outline the critical design requirements for emerging autonomous ICPS, starting with examples of key application drivers. We then explore the essential system specifications, examine how technological trends are shaping these advancements, and emphasize the importance of a holistic 3C co-design approach to effectively navigate the inherent complexities and trade-offs.

Across various industrial sectors, ICPS technologies are being extensively studied and evaluated to tackle critical challenges related to resource efficiency, real-time performance, and system resilience [7]. In process automation, for example, real-time wireless sensing and networking have demonstrated significant potential to enhance flexibility and improve real-time data collection in dynamic scenarios, such as the start-up phase of a paper mill or the adjustment of a chemical process

[8]. The advent of Industry 4.0 has lowered the barriers for implementing ICPS solutions within the manufacturing industry, where Cloud-Fog Automation is poised to make early and vital contributions [9]. The ongoing transformation of the energy grid, characterized by increased renewable electricity production and flexible consumption, presents substantial opportunities for developing energy management systems that systematically integrate intelligent *prosumers* (both producers and consumers) into the grid, including industrial plants and large buildings. In smart logistics and intelligent warehousing, the seamless integration of connected goods with mobile robotics and other automation technologies is catalyzing a transformative shift [10], with the potential to extend across broader segments of the supply chain. This includes the cyber-physical control of the entire freight transport system, integrating automated trucks and ships [11].

The design requirements for autonomous ICPS must effectively navigate several challenging trade-offs among control, communication, and computational capabilities. The core issue in networked control system design is on how to ensure a certain level of control performance despite limitations or uncertainties within the communication and computational platforms. Key research questions include how to design controllers that can manage network imperfections (such as delays, losses, and outages), how to tune communication protocols to meet the demands of real-time control loops, and how to jointly optimize control, communication performance, and resource utilization. Frameworks to address these challenges, particularly in wireless sensor/actuator networks, have been systematically developed and rigorously tested in practice [12]. More recent approaches involves goal-oriented integration of sensing, communication, computing and control [13], and regret-optimal cross-layer control designs.

Many automation and control systems operate under stringent timing requirements. The influence of time delays on closed-loop system performance is a well-established topic in control theory [14]. When considering ICPS, the underlying time architecture and its assumptions become even more crucial. The presence of time delays, jitter, and synchronization can significantly impact system performance [15]. In fact, control systems may lose stability if an adequate amount of data is not consistently circulated. It is possible to precisely determine the required data rate to maintain stability or achieve specific control performance [16]. Given that sensor or control data can be intermittently lost, estimation and control algorithms have been developed to address issues arising from lossy communication networks [17].

From a communication perspective, these foundational properties of networked control have given rise to the concept of semantic/goal-oriented communications, where the significance of specific data is considered in the network protocol [18]. The value of information contained in a data packet heavily depends on its location within the network and the timing of its delivery. Network resources should not be wasted on transmitting outdated data. The urgency of executing feedback control actions has been leveraged in event-triggered control, where sensing or actuation data are transmitted only when necessary. This approach has been

extended to various scenarios, such as event-triggered model predictive control. Finally, cyber-physical security and privacy requirements are naturally imposed on networked control systems, as the confidentiality, integrity, and availability of information across the feedback loop are paramount for many real-time and safety-critical control systems.

C. New-Generation Digital Infrastructure

An evolving computing and communications infrastructure is currently enabling a digital transformation in ICPS, creating a new-generation digital infrastructure. The foundation of new generation industrial systems builds on distributed computing and newer wired/wireless networking capabilities to support time-critical and other traffic types in a converged network. As computing resources at the edge/cloud take a key role in tasks that require high processing power, such as AI and other computational tasks that involve large amounts of data, the network (both wired and wireless) needs to guarantee lower latency with higher reliability. Avoiding congestion-induced delays and losses is paramount for such converged networks.

TSN technologies have emerged as a toolbox to enable converged networks where mixed criticality applications co-exist. TSN standards defined by the IEEE 802.1 Task Group include features for time synchronization, low latency, reliability, and resource management that can be enabled on top of Ethernet-based local area networks (LAN) [19]. These features can be combined to build a network infrastructure in which time-sensitive traffic receives priority treatment to guarantee latency bounds with high reliability. Recently, TSN features have also been extended to Wi-Fi [20] and enabled over 5G networks [21]. Recent advances in IEEE 802.11 standards, such as new scheduling capabilities in 802.11ax and multi-link operation (MLO) in 802.11be have enabled newer generation Wi-Fi devices and networks that can provide much more granular QoS with lower latency. Combined with TSN enhancements, Wi-Fi 6 solutions have been used to enable end-based robotics and other industrial automation use cases [22]. The latest Wi-Fi 7 chipsets based on the IEEE 802.11be amendment include the capability to operate in multiple links that can be used to not only reduce the channel access latency, but also to increase reliability through redundant transmissions and guarantee low latency performance during roaming [23].

5G standards have also introduced uRLLC capabilities that can be used to support high QoS guarantees for time-sensitive traffic streams. The 5G Release 16 specification has also defined mechanisms to integrate a 5G system with TSN-based networks and enable time synchronization and low latency guarantees in an extended wired-wireless TSN infrastructure. Recent research works on interfacing 5G capabilities with TSN have been demonstrated in [21].

Fig. 2 illustrates the envisioned computing and network infrastructure of the future that incorporate TSN and other latency and reliability enhancements in wireless protocols to support distributed computing. Computing resources for time-critical tasks may be available at distributed locations forming an edge computing infrastructure, which is served by a heterogeneous network including Ethernet, Wi-Fi and 5G technologies supporting TSN capabilities. The TSN features

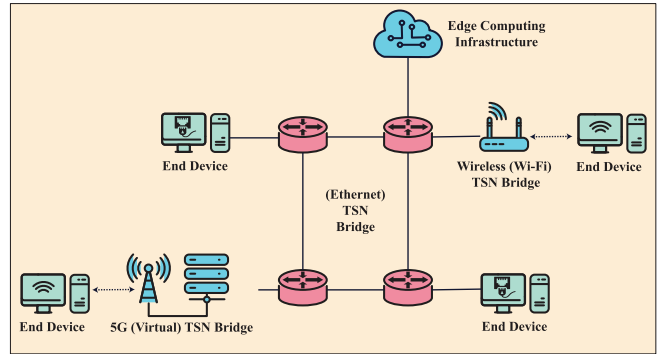


Fig. 2. Converged network architecture in TSN-based new-generation digital infrastructure.

can ensure accurate time distribution and guarantee latency bounds for time-critical flows across the network.

D. Closing the Gap With 3C Co-Design

Although there has been significant progress in innovating communication, computing and control technologies to support new ICPS applications and use cases, there is still a significant gap between the current/emerging digital infrastructure and the vision of an autonomous ICPS. The new-generation wired/wireless TSN-enabled digital infrastructure currently underway can only provide almost-deterministic communication at best, and does not effectively consider the adjacent computing and control requirements for achieving autonomous ICPS. Similarly, advancements in computing and control domains do not closely consider and tightly integrate with innovations in the underlying communication infrastructure.

To close this gap, we propose the adoption of Cloud-Fog Automation as the new paradigm to fully realize the vision of autonomous ICPS. In particular, we call for academic and industrial research communities to devote efforts into 3C co-design – co-designing technologies in communication, computing and control by considering the inter-related ICPS requirements. The central diagram in Fig. 1 visually captures the essence of 3C co-design. 3C co-design is a multi-dimensional synergy concept that is crucial in bridging the divide between current/emerging new-generation digital infrastructure and the realization of a fully autonomous ICPS. For instance, with the deployment of new wireless technologies to empower almost-deterministic uRLLC, a joint design of optimal control and computing techniques is non-negotiable for time-critical ICPS. It is also imperative that functional safety and security of such systems are not compromised. A prime focus will be on system-wide/application-level performance, driven by goal-oriented communication of interconnecting communication and networking technologies.

In the following sections, we further expand on the aforementioned research challenges and future directions by providing a detailed exploration of the three domains of 3C co-design: (i) goal-oriented communication, (ii) virtualization-empowered computing, and (iii) QoS-aware control. We subsequently address cyber-physical security as a significant

concern that intertwines with each of the 3C domains. Each section offers an overview of the respective area, reviews current and emerging technologies in the literature, and examines their inter-connections within the 3C domains. Each section outlines future directions respective to its own domain.

III. GOAL-ORIENTED COMMUNICATION

Historically, the development of mobile wireless communication systems has been driven by the increasing demand for the transmission of higher amounts of data by individual users, as well as the rapidly growing number of users. Therefore, the design and optimization of communication systems has been targeted at maximizing communication efficiency in terms of coverage costs and the delivery of information bits from a transmitter to a receiver. This is expressed in well-known communication metrics such as spectrum usage efficiency, network capacity, bit error rates, and packet error rates. Significant advances have been made in PHY and MAC layer techniques, such as channel encoding, hybrid automatic repeat request (HARQ) mechanisms, multiple-input multiple-output (MIMO) and beam-forming techniques, improved channel state information (CSI) estimation and signaling protocols, as well as resource allocation and scheduling algorithms. As a result, within this historical optimization paradigm, most of the additional network capacity is now primarily available in very high frequency bands, such as in millimeter bands (mmWave) and beyond.

In the development of the fifth generation (5G) mobile wireless technology, the focus of communication have expanded beyond human-oriented consumption to include machine-oriented communication. With the wireless communication industry's increased interests in IIoT and ICPS use cases, significant research and development efforts have been devoted to minimizing end-to-end latency, which is evident by 5G's uRLLC specification. However, it remains a challenge for wireless networks to address various applications needs and requirements, including bandwidth efficiency, low-latency, and low power/energy efficiency. Furthermore, communication technologies need to be adaptive to networks with very different communication demands, traffic and application patterns, which is often the case in private enterprise networks, and machine-oriented communication in ICPS.

In recent discussions about the future challenges in the development of wireless communication technology, in addition to further optimizations on PHY and MAC layers (for example with novel AI/ML approaches), there is a growing interest for a more holistic and application-driven optimization at the semantic level, which goes well beyond data bit or packet delivery. The pertinent problem of communicating the meaning of the data was first introduced by Shannon and Weaver in 1949 [24], where their work divides the broad subject of communication into three levels:

- Level A: How accurately can the symbols of communication be transmitted? (*The technical problem.*)
- Level B: How precisely do the transmitted symbols convey the desired meaning? (*The semantic problem.*)
- Level C: How effectively does the received meaning affect conduct in the desired way? (*The effectiveness problem.*)

Most of the works in communication and information theory, design and optimization since the introduction of Shannon's work have been focusing on addressing the *technical problem* on how to deliver bits and symbols of communication. With the introduction of the *semantic problem*, the primary goal of data communication is to accurately deliver the meaning of the data from a sender to a receiver. In many applications, this implies significant reduction of the data amount to be transmitted as semantically irrelevant data can be withheld from transmission. Specifically, when discussing machine-to-machine (M2M) control communication, the meaning of the information is extracted and processed by a control application. As a result, there are natural interests in designing communication mechanisms to progress to Level C and optimize the *effectiveness problem* of communication with the consideration of control goals of the application.

Recently, the subset of semantic communication which explicitly considers application goals or tasks is named as "goal-oriented communication". Fig. 3 presents the general system model for goal-oriented communication. Although it naturally extends semantic communication for machine-type control applications, goal-oriented communication is not limited to such use cases. It can be extensively applied to human-oriented communication. Some research works have started to go beyond communication and include aspects of compute, cache, and sensing jointly in alignment with control application goals [13]. We refer the readers to several excellent visionary, overview and survey articles on goal-oriented communication, which appeared in recent years [25], [26], [27], [28], [29], [30]. In the following sub-sections, we specifically highlight several research works and directions with the focus of goal-oriented communication relevant to the ICPS context.

A. Metrics of Goal-Oriented Communication

The first questions raised in the new paradigm of semantic and goal-oriented communication are on defining suitable metrics, such as how to formulate and measure the meaning of information, how to determine the communication goals, or how to define communication effectiveness. These key questions have been raised by the research community since the conceptualization of goal-oriented communication [30].

Age-of-Information (AoI) is one of the most important and novel metrics [31], [32]. As its name suggests, AoI indicates the time passed since the latest and not yet received data has been generated at the source (e.g., from a sensor). Although it does not have explicit semantic values, it indicates the freshness of information through its lifetime, which especially for ICPS applications often implies that fresher information is much more relevant and critical to act upon [33]. Higher AoI values translate to higher delays, where precise control and decision-making may not be possible. Time-critical ICPS applications can benefit from the integration of AoI at the semantic level to assist with task prioritization and resource scheduling. AoI has been already been used in several research

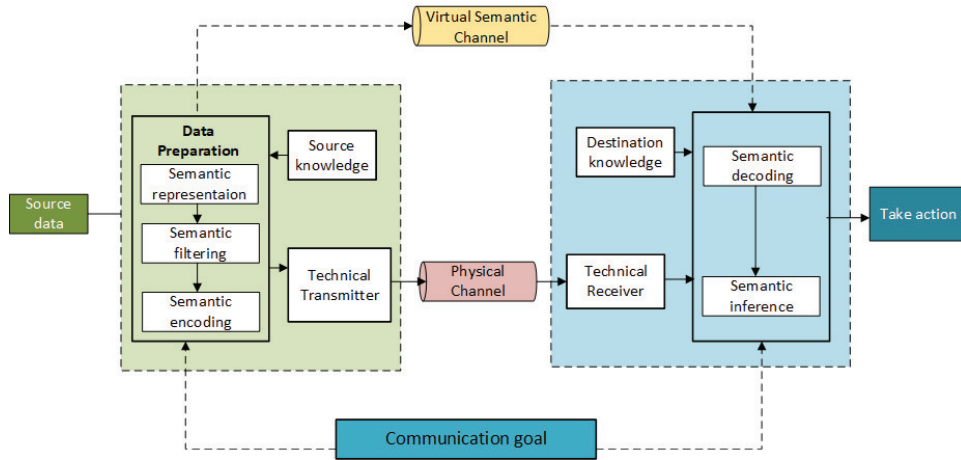


Fig. 3. System model for goal-oriented communication (adapted from [27] and [30]).

works, focusing on the co-design of control and communication systems in ICPS [34], [35].

Since the first introduction of AoI, there has been many adaptations to the AoI metric [30], [36]. In [37], the authors proposed *Value-of-Information* which extends AoI with a non-linear utility function to reflect the sensitivity of applications to the freshness of the information. With a similar motivation, the work in [38] introduced *Urgency of Information*, a novel metric which multiplies time-dependent context-aware weights and a cost function for the use of delayed information. The system urgency status for a new information update is characterized and used for data sampling at the source. It is also used for scheduling when multiple users share the same network resource. This intuitively makes sense, for example, when the system gets closer to certain critical threshold it is not allowed to cross, the urgency for the update increases. On the other hand, if the system is far away from a critical threshold, it is more delay-tolerant. The authors specifically investigated the performance benefits in a remotely control cart-pole systems (a mobile inverted pendulum setup), which accurately captures the control challenges in industrial control applications. In a similar manner, the authors of [39] introduced *Grade of Effectiveness* metric, which incorporates both freshness and usefulness attributes of the information.

An important innovation in the AoI metric for goal-oriented communication is the direction of information flow. AoI, as originally defined, is uni-directional by its nature and can be used either to optimize the downlink (DL) or uplink (UL) direction separately. In contrast, in control systems, the cycle from sensor to controller and then to the actuator needs to be taken into account as an end-to-end metric. The authors of [40] proposed the adapted *Age-of-Loop* metric to address this particular issue with AoI, and demonstrated the benefits in terms of control accuracy, where both DL and UL links are both jointly considered. In [41], the authors provided a good summary and comprehensive comparison of various AoI-adapted metrics in multi-directional information flow scenarios. They also introduced a more complex multi-dimensional tensor metric to unify multiple relevant aspects [42].

B. Goal-Oriented Communication Design

In the new goal-oriented communication paradigm (as presented in Fig. 3), nearly all aspects of communication can be re-visited and aligned with specific communication tasks with new optimization metrics.

1) *Data Preparation*: One of the initial important questions of goal-oriented communication is deciding which specific data needs to be send. Multiple goal-oriented and semantic optimizations can be undertaken at the source of the data itself, such as data quantization [43], data sampling [44], data compression [45], and source coding [46], [47]. With such measures, the goal-oriented communication system can better reduce the amount of data sent when the data is less critical for the communication goal; or increase if the controlled system becomes sensitive to sensor information and updates.

2) *Joint Source-Channel Coding*: Next, the question on how to transmit data in a goal-oriented and semantic manner needs to be addressed. With advances in deep learning, researchers have started to investigate joint source-channel coding – two processing tasks which are usually performed separately from each other due to their complexity. With the introduction of semantic communication, a joint integration of channel coding (technical transmitter in Fig. 3) and semantic encoding can be undertaken. Further, recent works have demonstrated that joint source-channel coding can be approached with deep learning techniques and it brings significant benefits for specialized applications (e.g., video transmission, image sensor data), which use joint data coding/protection that are aligned with wireless channel dynamics [48], [49], [50].

3) *Resource Allocation*: Due to the bandwidth constraints in wireless networks, the allocation of wireless resource becomes critical for reaching application goals. With recent advances in reinforcement learning (RL) algorithms, several research teams have devoted their efforts to apply RL methods to address the resource allocation problem with goal-oriented communication [50], [51], [52], [53]. These works include the use of multi-agent RL for multiple network users such as sensors and robots. It is demonstrated that such

learning algorithms are able to provide specialized action policies in complex dynamic network scenarios, where classical scheduling methods approach their limits. In [54], the authors extended the analysis to learning the effective allocation of communication and computing resources.

4) *Communication Protocols and Emergent Communication*: Finally, the protocol aspects for coordinating data transmissions from multiple agents can be optimized according to the design of goal-oriented communication systems. In [55], and [56], the authors have adopted an approach similar to multi-agent RL to investigate the optimization of pull and push-type protocols for certain control applications. In a more generalized approach, emergent communication among agents (which take actions), interact with environment and use communication to coordinate and cooperate so that their goal-oriented reward is maximized. This approach is investigated with RL frameworks in [57], [58], and [59]. These works represent early steps in this direction, and demonstrated the possibility and feasibility for the learning-based emergence of coordination and communication protocols for dedicated tasks such as industrial task offloading in ICPS [60].

C. Future Directions

While the rapidly evolving field of goal-oriented communication aims to overcome the limitations of current communication technologies, it remains in its early stages and faces many challenges and unresolved questions. Here we outline several key future research directions.

- 1) *Unified Theory*. Compared to conventional information theory, it is still challenging to derive a unified theory for semantic and goal-oriented communication to fully understand the fundamental limits and constraints of this new paradigm.
- 2) *Metrics*. There are many metrics discussed in the literature, which actively contributed to the design goal-oriented communication. However, these metrics are difficult to generalize, and they only address some aspects of goal-oriented communication aspects and limited to specific scenarios.
- 3) *Practical implementations*. To date, most of the research works in the field of goal-oriented communication are based on simulations or analytical frameworks. However, it is important to investigate the practical implementations of goal-oriented communication, and experimentally demonstrate the realized benefits in real hardware and wireless channels.
- 4) *Legacy co-existence*. Even when proved beneficial in some scenarios, new goal-oriented communication systems are likely to co-exist with conventional networks and share their resources [61]. The co-existence with legacy systems will bring about additional challenges in terms of performance benefits and trade-offs. Standardization efforts are also crucial in ensuring that these system designs are safe and practical for mass deployment.
- 5) *Beyond communication*. The paradigm of goal-oriented communication have started to expand beyond communication. New metrics can be used to take into

consideration other applications with relevant resources such as compute, caching, or wireless sensing. This trend has also motivated the need for an intrinsic 3C co-design in the context of ICPS, with a particular focus on industrial control applications.

Finally, it is imperative to further extend and clearly demonstrate the range of applications that could potentially benefit from goal-oriented communication. Currently, there is a particular focus on video, AR/VR and control-based applications. Further expansion, investigation and specification of various new and emerging applications are required from both academia and industry research.

IV. VIRTUALIZATION-EMPOWERED COMPUTING

Virtualization represents a fundamental computing principle, where the underlying hardware resources are logically partitioned/shared, and presented to the upper-layer software as if it had exclusive hardware access. It enables application scalability and the consolidation of multiple workloads over the same physical hardware [62], [63]. Since the off-the-shelf availability of hypervisors and virtual machines (VMs) two decades ago, virtualization has been rapidly adopted across various sectors. In particular, the cloud computing industry has widely adopted virtualization as its key enabling technology to drive the sharing/renting of computing resources to multiple clients. As the technology mature, virtualization has progressively proliferated ICPS and embedded system sectors, which possess strict functional safety and dependability requirements [64].

Innovations in ICPS and computing architectures such as fog/edge computing have demonstrated the feasibility of leveraging virtualization technologies in commercial off-the-shelf multi-core processors in industrial environments with the goals of achieving system-wide scalability, reliability, availability, security, durability, portability, and integration with legacy systems [65], [66]. However, despite much efforts in academic and industrial research, there are still significant challenges in adapting information technology (IT)-specific virtualization mechanisms towards hardware/software stacks in operational technology (OT)-based ICPS applications.

Fig. 4 illustrates the role of virtualization technologies in ICPS, which can be delineated into three distinct yet interconnected layers: (i) *application*; (ii) *communication*; and (iii) *hardware* virtualization. These virtualization layers span from the field devices, field network, fog/edge platforms, remote network, to the cloud platform. These inserted virtualization layers enable the “lower-layer independency”, which protects the business interests of the vendors, however, they also incur some “black holes” of overheads.

The notion of ‘*mixed-criticality*’ in virtualization is paramount in ICPS [67]. In cases where tasks with different safety-criticality are executed on separate VMs, a high level of temporal and spatial isolation between VMs are required to guarantee their safety and security. In cases that require dynamic reconfiguration of networks and migration of tasks across virtualized computing nodes, the safety of VM migration needs to be ensured so that the timing and execution behaviors are not adversely affected [68].

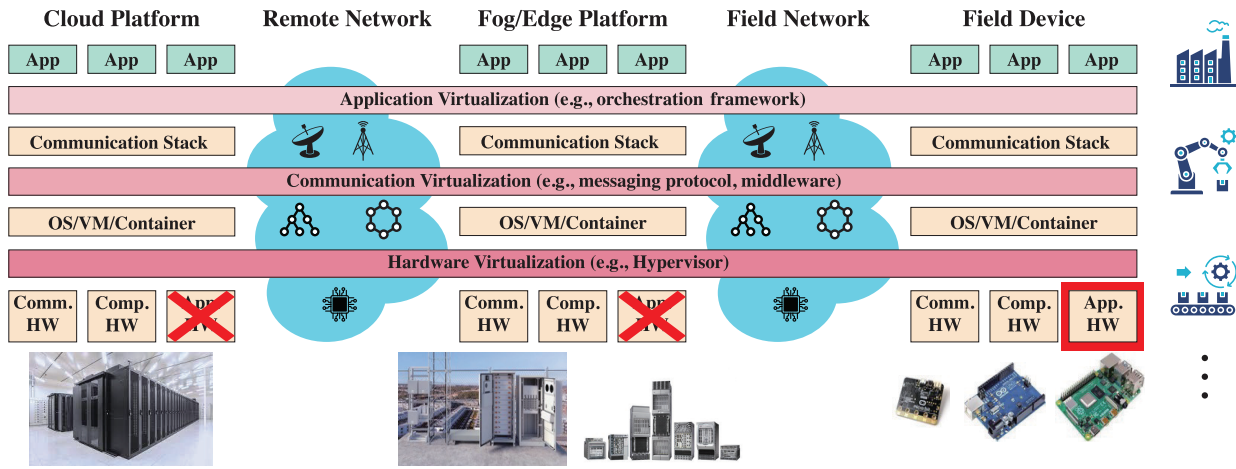


Fig. 4. Virtualized-empowered computing across application, communication, and hardware layers for executing control logics within autonomous ICPS.

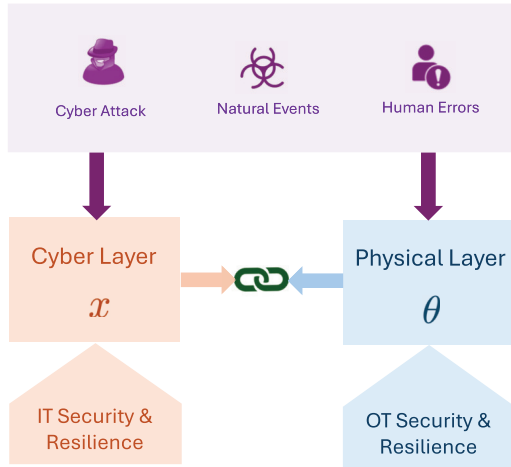


Fig. 5. An industrial control system representing the convergence of its cyber and physical layers. The state of the well-being of the IT system x and the state of the OT system θ are interdependent. They are exposed to a broad attack surface, including DDoS attacks, ransomware, malware, jamming, MitM attacks, and human vulnerabilities. Security and resilience mechanisms must be designed at each layer to account for the cross-layer interactions and inter-dependencies. A co-design approach is essential to improve the cyber-physical security of the IT and OT-convergent ICPS.

The aforementioned stringent industrial requirements led to the development of the concept of ‘deterministic virtualization’ [69], which specifies: (i) *determinism* in guaranteeing predicible latencies and low jitter; (ii) *safety certification* of VMs that run safety-critical control tasks; (iii) *isolation* of VMs to ensure system security and safety, and reduce non-deterministic timing originating from shared resources; and (iv) *flexibility* in replacing VMs and task sets during runtime for scalability, re-configurability and safe live migration.

While virtualization technologies have significantly matured and widely adopted at the hardware layer, it has recently found its way into the application and communication layers, particularly in ICPS. Hence, their tight interlocking and interactions with hardware virtualization for achieving deterministic virtualization merits the community’s attention. Below we present

a concise review of the state-of-the-art in ICPS-specific application, communication, and hardware virtualization.

A. Application Virtualization

1) *OS-Level Virtualization*: In contrast to the high performance overhead incurred by hardware virtualization (where the guest OS is executed inside a VM), OS-level virtualization using containers are much lighter weight and could promisingly address the stringent footprint requirements of real-time embedded systems in ICPS [70]. The key innovation of containers is the abstraction of OS processes, isolating the user-space while sharing the host OS kernel (e.g., Linux uses the *namespace* process isolation mechanism, and *cgroups* to provide resource management capabilities). Containers are virtual environments with local APIs, libraries and settings, with their own perception of virtual resources (e.g. CPU, memory, file systems, network interfaces, etc). The container start-up/shut-down times are minimal, and it can be built from templates and shared/extended/reused across different containers. Entire applications could be containerized and spread across different containers, making containers a versatile and light-weight solution for dynamic environments such as task scheduling and resource allocation in ICPS [71].

The adoption of OS-level virtualization in ICPS is a recent trend. The aim is to leverage containers instead of VMs to achieve isolation with minimal overhead while meeting the performance requirements of mixed-criticality systems [72]. Experimental studies and analysis have demonstrated the feasibility and benefits of using container and its in-built orchestration frameworks (e.g., Docker Swarm) and platforms (e.g., Kubernetes) in ICPS [70]. For example, the work in [73] presented an architecture for a container-based multi-purpose industrial controller, including support for real-time PLC execution engines and emulation of legacy engines. The work in [74] evaluated the time delay impacts of the Docker software with a real-world use case, and demonstrated that Docker containers can be used as a distributed lightweight virtualization scheme to meet the soft real-time requirements of industrial automation. In [75], the authors evaluated the performance of container-based solutions for

service provisioning in wired/wireless industrial clusters for improved manageability and scalability. A recent work in [76] proposed a containerized application-driven service architecture for digital twin networks using Kubernetes in industrial networked environments, and provided empirical application insights for a practical extension of the ITU-T Y.3090 standard.

2) *Microservices*: Microservices is a software development approach that structures an application as a collection of loosely coupled and independently deployable services [77]. Each microservice fulfills a specific business requirement, encapsulating its own logic, data, and communication interfaces. This architecture enhances system modularity, scalability, and fault isolation. Consequently, it enables teams to develop, deploy, and scale different parts of the system independently. Microservices typically communicate via lightweight protocols such as HTTP or messaging queues, and often deployed in tandem with containerization technologies. Microservices are suitable for complex, distributed systems where agility, continuous delivery, and resilience are critical. Building upon its success in the computing industry, research and development into ICPS-specific microservices and their orchestration frameworks are emerging, creating new opportunities for innovations and advancements [78].

IEC 61499 is a notable standard for system-level microservice orchestration. The work in [79] proposed a microservices-enabled orchestration architecture for deploying features specified in IEC 61499, with containers to provide maximum flexibility, interoperability, and efficiency for industrial edge applications. It maps the rules between IEC 61499 function blocks and microservice architecture for both the design time and runtime. In [80], the authors proposed a cyber-physical microservices framework for manufacturing based on unified modeling language (UML) models. Guidelines for functional requirement decomposition are formulated in [81] and modeled across the process, machine, and functionality layers of the process control systems. A Kubernetes microservices approach is proposed in [82] to encapsulate containers, with a special web transfer service for exchanging data between virtual PLCs (vPLCs). A cloud-based IEC 61499 virtual commissioning method is presented in [83] to orchestrate a set of application runtimes with microservices.

B. Communication Virtualization

1) *Software-Defined Networking (SDN) and Network Function Virtualization (NFV)*: Communication virtualization is primarily driven by the virtualization of the network layer, with technologies such as SDN and NFV. SDN separates the control plane from the data plane, which were deemed inseparable in proprietary hardware/software in the past. The SDN control plane makes traffic-handling decisions, while the data plane solely moves data content between two network nodes. SDN enables a global view of the network, and achieves hardware independence while providing intelligent policing/routing decisions, and ensuring redundancy and fault tolerance. NFV [84] is a resulting architecture that specifies the executing of SDN functions independent of any specific hardware platform. It is an infrastructure that administers and orchestrates virtual network functions (VNF), which are

individual network services running as a software-only VM instances on generic hardware [85].

SDN/NFV have been deployed extensively in data center networks but have only recently realized in closed and controlled ICPS networks [86]. Recent research efforts in designing/deploying centralized [87] and distributed SDN controllers [88] for network automation (including sensing/actuating, process orchestration/automation, etc) have demonstrated the applicability of SDN/NFV in ICPS [89]. However, the timely synchronization of resource demand information of VNFs with the SDN controller is critical as the delays directly impact the QoS of emerging applications (e.g., industrial augmented reality, digital twins, human-robot collaboration) relying on such information. Hence, the authors in [90] proposed a SDN/NFV network slicing architecture, with ML-based algorithms for efficient VNF mapping and scheduling. To address the challenges of VNF placement, the work in [91] formulated VNF placements as one-shot/NP-hard problems and proposed a heuristic online algorithm, whereas optimization modulo theories (MaxSMT) with expressive constraints and formal verification are proposed in [92].

Several recent works have focused on building experimental platforms that implements joint radio resource allocation and beam-forming with SDN/NFV [93], supports tactile Internet with novel 5G/6G uRLLC and blockchain-based methodologies in smart manufacturing [94], [95], and enables a resilient network with a pseudo-honeypot strategy [96]. Network slicing management with remote adaption and orchestration [97], transfer learning [98] and graph neural networks [99] are also recent research trends.

2) *Virtual Network Embedding (VNE)*: Network virtualization technologies such as NFV rely on algorithms that can instantiate virtualized networks on the infrastructure to optimize the layout of service-specific metrics/indicators [100]. Algorithms that dictate these behavioral properties are referred to as VNE. VNE algorithms solve virtual network requests and efficiently allocate physical node/link resources. Since VNE attempts to optimize for multiple constraints, including attributes such as CPU, storage, bandwidth, request admissions, and other dynamic requirements, it is classified as an NP-hard problem and could benefit from AI/ML-based strategies [101].

VNE algorithms governs the elasticity of virtualized node and link resources in ICPS networks, and it is critical that VNE satisfy the multi-dimensional QoS requirements. For instance, the work in [102] proposed a dynamic embedding scheme to optimize performance of each service node, and as each node satisfies the QoS requirements, a re-embedding scheme of the heuristic algorithm will be performed. This scheme ensures the flexibility of virtual nodes and links assignment while fulfilling the customized QoS requirements. In [103], the authors proposed an intelligent latency-aware VNE scheme to provide deadline guarantees for industrial virtual networks by implementing both static embedding (with an any-path algorithm) and dynamic forwarding (with deep-learning). Application-driven VNE approaches with an any-path routing scheme are proposed in [104]. These approaches exploit unique features in wireless channels in factory automation systems.

C. Hardware Virtualization

1) *Virtualized PLC (vPLC)*: vPLC is introduced to address the limitations and challenges posed by physical PLCs [105]. vPLC is a hardware-agnostic abstraction of the control unit and memory functions of a PLC, however, it still requires an interface to communicate with the I/O modules. vPLC is able to incorporate multiple control functions and provide benefits in terms of processing capabilities, flexibility and efficient resource use interoperability and optimization between different I/O module vendors, reducing device density in factory floor, and cost savings. It integrates both IT and OT software components, and plays a critical role towards an integrated and autonomous ICPS [106].

A recent work in [107] presented a methodology for comprehensively evaluating the communication performance of vPLCs (based on a set of defined metrics from corresponding transaction round trip time, and the transmission latencies between devices) when exchanging data for supervision, coordination, and control with other machines and supervisory control and data acquisition (SCADA) in machine-to-machine (M2M) scenarios. A comparison platform is designed to compare against real PLCs, along with the creation and validation of an analytical model with full network traffic access, which can be used for simulators for worst-case analyses. Experiments indicated that vPLC could work as fast as a real PLC with minimum communication latency, but could potentially induced randomized delays (similar to jitter) due to the Internet Protocol (IP) stack implementation of the vPLC.

2) *Real-Time Hypervisors*: Real-time hypervisors add explicit support for managing the time budget allocated to VMs with specific scheduling algorithms, to assure that individual VMs comply with the stringent and explicit timing constraints [108]. Dynamic hypervisors map VM-related resources as needed, whereas static hypervisors provides a one-to-one mapping between virtual and physical hardware resources. Static solutions have lower performance overhead and they are more robust to failures and mis-configurations, hence they are often embedded in safety-critical and mixed-criticality applications [67]. The work in [109] proposed a dynamic VM migration approach with a novel online scheduling algorithm based on minimal distance tree - heuristic breadth first search (MDTC-HB) construction, in multi-cast TSN networks to guarantee the firm latency of control messages. Experiments demonstrated the efficacy of the framework in terms of rapid response time while only consuming 50% bandwidth. In [110], the authors proposed a virtualized platform that specifically address the hard real-time requirements of advanced control mechatronic systems. The platform uses the Jailhouse hypervisor and designates micro-controllers with specific/routine computational tasks, and offloads heavy/hard real-time computations to micro-processor cores, which is analogous to the recommendations provided in [111].

D. Future Directions

There are many exciting future research directions in virtualization-empowered computing to support 3C co-design

towards autonomous ICPS. We present several promising directions below.

1) *Application Virtualization to Support TSN*: Although advances have been made in designing deterministic execution of control and computing tasks in virtualized environments, typical use cases are still self-contained in static contexts, and originates from cloud computing [69]. Therefore, future work will need to rethink a new paradigm for real-time hypervisors, containers and their orchestration frameworks to support deterministic virtualization with a particular focus on highly granular time-sensitive ICPS networks such as TSN, which includes implementing a deterministic I/O, dynamic, coordinated VM, task and network schedules for flexible yet safe control in virtualized platforms.

2) *Knowledge-Centric Networking (KCN) With SDN/NFV*: KCN is a new and exciting field of research that emerged from content/information-centric networking (CCN/ICN), which leverages in-network computing, in-network storage, and in-network communication to create the needed knowledge, which allows sensing devices to obtain useful data from anywhere in the network [112]. Much of the related work are performed in the public Internet networks, and research efforts on KCN in ICPS networks are limited. With many communication virtualization technologies already driven by SDN/NFV, KCN is a new kid-on-the-block that could significantly improve network analytics and intelligence in ICPS, providing a complete framework from front-end sensing to back-end networking.

3) *AI/ML for Virtualized System Testing and Certifications*: Industry standard compliance is paramount when deploying virtualization in ICPS. The wide variety of hardware virtualization approaches and their complexity present a challenge in ensuring standard compliance [67]. For example, emerging approaches such as micro-kernels provide strong security and effective verification, but still yet to be fully safety-guaranteed when deploying in industrial environments. There is also a lack of benchmarks and effective test suites to produce evidence in the safety certification process. Therefore, future research could focus on leveraging AI/ML to ensure the compliance of safety and security guidelines through rigorous robustness and performance testing processes, while also informing future development and evolution of industry standards.

V. QoS-AWARE CONTROL

QoS-aware control is a critical component in the design and operation of ICPS, where precision, reliability, and timeliness are fundamental to the success of autonomous industrial operations. In ICPS, the convergence of industrial control and computational systems necessitates a tight coupling between communication networks and control mechanisms. These systems rely on the ability to manage diverse processes in real time, ranging from machinery control to sensor data analysis, hence requiring sophisticated control strategies to ensure system stability and performance under strict QoS requirements. QoS-aware control mechanisms are designed to guarantee these requirements, ensuring that factors such as bandwidth, latency, reliability, and availability are optimized

to support time-sensitive and mission-critical industrial processes. Without proper QoS management, the dynamic and complex nature of ICPS can lead to performance bottlenecks, performance degradation, and even system failures that could dramatically disrupt production and compromise safety [113].

In the Cloud-Fog Automation paradigm, QoS-aware control is critical for achieving determinism in communication, computation, and control. Since Cloud-Fog Automation is a network-centric architecture where cloud and fog resources collaboratively handle the computational and control requirements of ICPS, it creates a distributed yet synchronized ecosystem. In this architecture, fog nodes located closer to the physical processes in the industrial environment can handle tasks that require low latency and high reliability, while the cloud offers scalable computational power for data-heavy tasks such as predictive maintenance, system optimization, and large-scale data analytics [114]. By leveraging this hybrid cloud-fog environment, QoS-aware control ensures that computational resources are efficiently allocated and that control actions are performed within the strict time constraints required by industrial operations. This approach minimizes latency, optimizes bandwidth usage, and ensures real-time decision-making, which are essential for maintaining the operational integrity of autonomous ICPS [115].

A key challenge in QoS-aware control within ICPS is the need to manage dynamic, unpredictable workloads while ensuring the system meets pre-defined performance thresholds. Traditional control systems, which are often built on static models, struggle to adapt to the dynamic nature of industrial processes where workload fluctuations, network variability, and potential system failures are commonly observed. In contrast, modern QoS-aware control mechanisms employ adaptive strategies that continuously monitor system performance and adjust resources in real time. These systems use feedback control loops, where control decisions are based on real-time data from sensors and actuators across the entire ICPS. Techniques such as model predictive control (MPC) are used to anticipate system behavior and adjust control actions accordingly to prevent QoS violations [116]. Furthermore, ML models are increasingly integrated into these control frameworks, enabling predictive QoS management where potential performance degradation is identified before it occurs. This predictive capability is especially valuable in ICPS, where system downtime or delays can lead to costly disruptions in industrial processes [117].

Security and reliability of QoS-aware control in ICPS are growing concerns as industrial systems become more connected and complex. Cyber threats and system vulnerabilities can have disastrous consequences in ICPS, making security-aware QoS management a critical area of research. Emerging technologies such as blockchain are being explored for their potential to enhance the transparency and integrity of QoS data in industrial environments, providing an additional layer of protection against data tampering and unauthorized access [114]. Additionally, advanced communication technologies such as 5G are expected to play a significant role in enhancing QoS by providing ultra-reliable low-latency communication channels that are crucial for real-time control. By combin-

ing these technologies with sophisticated QoS-aware control mechanisms, the future vision of autonomous ICPS can be realized with greater reliability, security, and efficiency, and paves the way for new industrial applications [115].

A. State-of-the-Art of QoS-Aware Control

QoS-aware control is vital for achieving the high-performance requirements of autonomous industrial operations. The control mechanisms in these systems must handle complex, real-time interactions between machines, sensors, and actuators, while accounting for communication delays and computational constraints. Unlike traditional control systems, QoS-aware control takes into account not only the control objectives but also the communication and computational infrastructure that supports it. Within this framework, latency-aware control has emerged as a prominent approach to manage time-sensitive processes. Latency is often the most critical QoS metric in ICPS. Delays in feedback or actuation can cause performance degradation or even system failures. Latency-aware control strategies, such as time-triggered control (TTC) and event-triggered control (ETC). TTC and ETC are particularly effective in managing communication-induced delays by synchronizing control actions with predictable communication intervals. TTC schedules control updates at fixed times, ensuring that control signals are always sent within a pre-defined time frame, whereas ETC dynamically adjusts control actions based on the system's state, hence reducing the communication load when the system is stable and increasing the load during critical changes.

The integration of communication and control is essential for managing the strict timing and computational demands of industrial processes. MPC is one of the most widely used methods in this context, and has been adapted for latency-aware environments [118]. In latency-aware MPC, control actions are calculated not only based on the current state of the system but also with an anticipation of potential communication delays and computational constraints. This anticipatory control framework allows the system to maintain optimal performance even when data transmission or computation takes longer than expected. In addition, latency-aware MPC can be combined with resource management techniques, where computational resources are dynamically allocated to processes that are most sensitive to delays, thus ensuring that critical operations are prioritized [119]. This approach is often implemented in cloud-fog architectures, where the fog layer handles latency-sensitive tasks closer to the operational environment, while the cloud is responsible for less time-critical, computationally intensive processes. This distribution of computational and control tasks between cloud and fog layers ensures a balanced workload, and reduces the risk of latency-related QoS violations in ICPS.

Another key aspect of QoS-aware control is the co-design of communication and control strategies [120], which addresses the direct interaction between communication delays and control performance. In traditional control systems, communication and control are often treated independently, leading to sub-optimal performance in networked environments. However, control strategies are designed to explicitly account

for the underlying network's behavior in QoS-aware ICPS, including variable communication delays, packet loss, and bandwidth limitations. A typical example is the networked control systems (NCS) framework, where control decisions are made based on the available communication bandwidth and network conditions. In this context, ETC is particularly effective as it reduces unnecessary communication by only sending updates when a significant change in the system state occurs [121]. ETC-based systems is able to dramatically reduce communication overhead while maintaining high control accuracy, making them an ideal candidate for industrial processes that require tight control but operate over bandwidth-limited networks. These techniques ensure that control systems are not overloaded with frequent updates, which could introduce unnecessary delays, but instead are triggered by meaningful events that impact the process.

Furthermore, the application of adaptive and ML-driven control in latency-sensitive environments has become a critical focus in the evolution of QoS-aware control for ICPS [122]. Adaptive control methods are increasingly being employed to dynamically adjust control actions based on changing system states and varying network conditions. For example, latency-aware adaptive control (LAAC) continuously monitors the network and computational latency to adjust control actions in real time, allowing the system to adapt to fluctuating conditions [123]. ML techniques have been integrated into ICPS to enable systems to learn optimal control strategies over time, particularly in environments where the relationship between communication delays and control actions is too complex to model with traditional methods. These learning-based models can optimize control strategies dynamically by observing how different control actions affect system performance under various latency conditions, thus enhancing both responsiveness and efficiency. For instance, RL-based methods can be effectively applied to automatically adjust control update rates in response to network conditions, hence ensuring that communication resources are used efficiently while adhering to strict QoS requirements.

B. QoS-Aware Communication and Control Co-Design

Through the lens of 3C co-design, QoS-aware control aims to integrate communication, computing, and control in a unified manner to achieve optimal performance in complex systems [124]. One key technique is the joint optimization of communication and control systems. This method ensures that control systems are aware of communication delays and packet losses, while communication protocols are designed to prioritize the timely delivery of control signals. Research shows that co-design frameworks such as MPC combined with NCS can help achieve this balance. This joint optimization framework ensures that neither the control objectives nor the communication performance is compromised, which is essential for real-time applications in autonomous ICPS.

Another crucial aspect is the dynamic adaptation of control strategies based on real-time feedback from the communication network [123]. Control algorithms should adjust their operations depending on the current state of the network, such as adjusting control update rates based on available bandwidth

or reducing the control task frequency when network delays occur. Techniques such as adaptive control systems or ETC have proven to be effective in adjusting the frequency of control commands based on network conditions. This adaptability ensures that the system maintains stability and performs optimally, even in constrained systems. The ability to react to varying QoS metrics such as latency, jitter, and packet loss in real time is critical for control systems deployed under the Cloud-Fog Automation paradigm, where network conditions are not always predictable due to industrial services being deployed in a distributed fashion.

Efficient resource allocation also plays a vital role in ensuring that QoS-aware control systems operate effectively [5]. In distributed control systems, resources such as bandwidth and computational power are limited. Therefore, resource management techniques that prioritize critical control actions over less important data transmissions are essential. Priority-based resource allocation methods, such as those that use optimization algorithms to assign bandwidth to the most time-sensitive control signals, have been shown to improve system performance [125]. In networked industrial systems, where communication congestion can significantly impact control accuracy, effective resource management ensures that communication delays are minimized, allowing control loops to function correctly. This dynamic allocation is especially important in large-scale systems, where many devices must share limited communication and computational resources.

Finally, real-time feedback loops and cross-layer designs are crucial for ensuring optimal performance in QoS-aware control systems. Real-time feedback allows control systems to adjust their operation based on current communication performance, hence ensuring responsiveness to changing network conditions. Cross-layer design further enhances the system by integrating control requirements directly into the communication protocols, making the communication system more responsive to control demands. Research in cross-layer optimization techniques, which includes integrating physical and network layers, demonstrates that aligning control and communication objectives leads to more efficient operation in large-scale ICPS. The interplay between control, communication, and computing ensures that all three domains operate in harmony, leading to enhanced performance and stability in future autonomous ICPS.

C. Future Directions

There are currently several key research directions within the QoS-aware control domain, which will form critical enablers for 3C co-design in autonomous ICPS.

1) *Adaptive Control Under Communication Variability:* Autonomous ICPS will heavily depend on advanced QoS-aware control strategies that ensure seamless interaction between control, communication, and computing. The primary challenge for QoS-aware control lies in ensuring that control systems can function optimally despite the inherent variations in communication and computing capabilities across cloud and fog environments. Control algorithms need to be adaptive and intelligent. They need to respond to changes in communication delays, computational loads, and system dynamics in real time

[126]. The challenge is that QoS-aware control must manage the timing and quality of the information exchange between distributed systems while also ensuring that control actions meet the strict performance criteria required for industrial automation.

2) *Real-Time Control Coordination With Communication and Computing*: QoS-aware control not only maintains the stability and precision in control loops but also ensures that control decisions are made in coordination with communication latency and computing resources [127]. As the cloud-fog infrastructure handles distributed tasks, the real-time nature of control actions becomes even more critical. Latency and computation delays must be accounted for in control strategies, ensuring that key decisions are made within the timing constraints of the system. For example, real-time control systems operating in environments such as manufacturing or process automation need to adjust dynamically to fluctuations in network performance and computing availability, while simultaneously avoiding system slowdowns or inefficiencies.

3) *Scaling QoS-Aware Control for Large-Scale ICPS*: The interplay between control and communication becomes especially important when ICPS need to scale, which involves a large number of interconnected devices and sensors [128]. QoS-aware control must leverage distributed control architectures to scale, where control decisions are decentralized and localized to reduce communication delays. However, the co-design of control with communication protocols ensures that vital information is still shared across the system in a timely manner, without overwhelming the network. For instance, control algorithms must prioritize critical data flows while ensuring non-essential data does not interfere with time-sensitive control decisions. This coordination allows industrial systems to operate smoothly, thus maintaining stable system performance and avoiding bottlenecks due to communication constraints.

4) *Control Algorithms Co-Designed With Computing Resources*: As control systems increasingly rely on data-intensive decision-making, computing resources play an important role in QoS-aware control. With edge/fog computing, the computing load can be distributed to process control tasks closer to where they are needed, hence reducing the dependency on cloud resources [129]. This mechanism reduces latency and allows control systems to act faster, and enhances the real-time responsiveness of ICPS. Autonomous ICPS will require control algorithms that are tightly coupled with computing systems, and also dynamically adjust based on the availability of distributed computing resources. Through 3C co-design, future QoS-aware control systems will achieve higher levels of efficiency, scalability, and adaptability, hence meeting the increasingly complex demands of industrial automation and real-time operations.

VI. CYBER-PHYSICAL SECURITY

Cyber-physical security focuses on protecting industrial control systems from cyber threats that can disrupt physical processes. In the context of autonomous ICPS, this security aspect becomes even more critical due to the increased connectivity, integration, and convergence of IT/OT. Ensuring the

security of these systems involves protecting the integrity, availability, and confidentiality of both cyber and physical components. Cyber-physical security strategies must address threats that can affect everything from data networks to physical machinery, potentially causing significant operational disruptions, safety hazards, and financial losses.

The need for cyber-physical security co-design arises from the inherently intertwined nature of ICPS, which is divided into distinct yet interdependent cyber and physical layers. The *cyber layer* comprises assets such as servers, workstations, sensors, communication networks, and networking components. In contrast, the *physical layer* includes actuators, controllers, and factory plants. This tight coupling requires an integrated security approach that considers the mutual dependencies and interactions between these layers. One example of this tight coupling is in large-scale power networks. When a cyber attack targets the communication network, it could disrupt the data that is being transmitted from the sensors to the control systems. This disruption might lead to incorrect readings or commands being sent to actuators and controllers. As a result, an attacker could potentially cause power outages, equipment damage, or even trigger cascading failures across the network [130].

Historically, many legacy systems in critical infrastructures are isolated from networks, making it difficult for attackers to access sensitive physical system information. However, the integration of ICPS with cloud computing and other modern information technologies has exposed these systems to a wide range of cyber attacks. The adoption of cloud computing in ICPS has brought numerous benefits, such as enhanced scalability, improved data analytics, and cost efficiency. Cloud computing enables real-time monitoring, remote control, and data storage, which are essential to optimize industrial processes and improve operational efficiency. Despite these advantages, cloud computing also introduces new vulnerabilities and attack vectors [131]. Traditional cloud security solutions, which focus primarily on data confidentiality and integrity, are insufficient for ICPS protection due to the critical need for data availability and real-time data for system stability. ICPS operations are highly dependent on timely and accurate flow of information to control and monitor physical processes. Any delay or interruption can have severe operational and safety implications. For instance, latency issues in cloud-based communications can disrupt the synchronization of control systems, leading to potential malfunctions or safety hazards.

Modern attackers employ sophisticated methods, such as advanced persistent threats (APTs), which can infiltrate systems over extended periods, compromising secret keys and sensitive information before executing further stages of an attack. Additionally, the inter-dependencies between the cyber and physical layers present further challenges. Effective ICPS security requires a holistic approach where both layers are co-designed. The cyber layer must support the physical layer's performance demands, while the physical layer must ensure that the cyber layer's security measures are robust and resilient. This integrated design improves the overall security and resiliency of ICPS. To address these challenges,

co-designed security frameworks that consider the unique requirements of ICPS environments are pertinent. These frameworks should integrate advanced threat detection and response capabilities, leveraging AI/ML and data analytics to identify and mitigate potential threats in real time.

A. *Cyber-Physical Security Threats*

Cyber-physical security issues in cloud-enabled industrial automation systems present significant challenges due to the complex and interconnected nature of these environments. Industrial control systems, including SCADA systems and distributed control systems (DCS), are integral components of industrial automation. These systems are increasingly integrated with cloud computing and fog computing architectures to enhance efficiency, scalability, and real-time decision-making. However, this integration also introduces new vulnerabilities and attack vectors. Key threats include Denial-of-Service (DoS) and Distributed Denial of Service (DDoS) attacks, which disrupt system availability by overwhelming resources, and Man-in-the-Middle (MitM) attacks, where attackers intercept and manipulate communications, leading to unauthorized access and compromised system responses. In addition, APTs pose complex, prolonged risks that exploit vulnerabilities over time to steal data or disrupt operations. Ransomware attacks target ICPS' dependency on data availability, encrypting critical information, and demanding payments for decryption, while supply chain attacks compromise systems indirectly by exploiting vulnerabilities in the production or delivery chain of hardware and software. Finally, physical security breaches allow attackers to manipulate or disrupt systems through direct access to components such as servers and control hardware, often complementing cyber attacks to amplify their impact.

Addressing these threats requires a comprehensive, integrated security approach. This includes robust defenses such as real-time monitoring, advanced detection mechanisms, and co-design strategies that combine physical and cyber security measures. Innovations such as blockchain for supply chain integrity, zero-trust architectures, and multi-layered models are being explored to enhance resilience. A unified understanding of the interconnected nature of these threats is critical for developing strategies that mitigate risks while maintaining system performance, ensuring the integrity of ICPS against an evolving and complex threat landscape.

B. *State-of-the-Art of Cyber-Physical Security Co-Design*

The co-design of cyber-physical security solutions requires impact-aware strategies that account for how security measures affect overall system performance, such as communication delays or signal-to-noise ratio (SNR). For example, the work in [132] explores impact-aware power control mechanisms to optimize network control systems alongside wireless communication networks. This approach improves both system efficiency and resilience against adversarial behaviors, highlighting the importance of designing security measures that balance robust protection with minimal impact on overall system performance.

In addition to ensuring that cyber designs are impact aware, it is equally important to design physical control systems taking into account the performance limitations imposed by the cyber layer. The effective integration of these layers is crucial to optimizing the overall performance of the system. For example, communication-aware control design is a key approach for developing networked control systems. This method involves designing control strategies that account for communication constraints such as delays and packet drop rates, which can significantly impact control performance. Several studies have contributed to this field, including [133], [134], [135], which explore optimal control performance by integrating communication parameters into the control design.

C. *Design for Resilience*

In cloud-enabled industrial automation systems, traditional security measures must evolve to focus on resilience, as perfect cyber-physical security is unattainable and vulnerabilities are inevitable. Resilience-oriented design emphasizes building systems that can withstand and recover from cyber-attacks with minimal operational disruption [136]. For example, recent research highlights mechanisms to quantify and mitigate risks in cloud-enabled networked control systems (CE-NCS) [137], which often outsource heavy computations to cloud servers. These systems face challenges in ensuring trustworthiness in their cyber-physical connections and cloud dependencies, necessitating robust strategies to maintain operational integrity.

Redundancy is a key resilience strategy that involves backup systems and duplicate network paths to ensure continuity even when parts of the system are compromised. Studies on IoT networks and the Internet of Battlefield Things (IoBT) have demonstrated the value of adaptive network designs that can reconfigure dynamically in response to disruptions. These designs improve robustness against threats such as DoS and jamming by incorporating optimal strategies that adapt to changing network topologies, ensure continuous operations, and safeguard critical processes.

Self-healing adaptive systems further bolster resilience by detecting, isolating, and repairing damage from cyber-attacks in real time. These systems use automated response mechanisms to maintain functionality without human intervention, adapting to evolving threats and minimizing disruption. By integrating these resilience-by-design frameworks with traditional cybersecurity measures, organizations can create robust, adaptive defenses that maintain functionality and recovery capabilities in the face of persistent and evolving cyber threats. This holistic approach is essential for the security of complex industrial environments.

D. *Future Directions*

Recent rapid advances in AI and 5G communication networks have introduced a wide range of promising research directions in cyber-physical security and system resilience. As these technologies evolve, they offer both unprecedented opportunities and new challenges for securing complex and interconnected systems. The integration of 5G technology into ICPS transforms their operational capabilities, allowing for

more efficient, real-time communication and data processing. However, this also expands their attack surface, making them more vulnerable to cyber threats.

1) *Research Directions Enabled by 5G*: The enhanced connectivity offered by 5G means that ICPS components, such as sensors, controllers, and actuators, are more interconnected than ever before. While this promotes greater automation and efficiency, it also introduces new entry points for attackers. As more devices and systems communicate with the cloud and each other, the number of vulnerable nodes increases, creating more opportunities for attackers to exploit weak points in the network. The larger and more complex the system becomes, the greater the security challenges.

As 5G enables real-time data processing in the cloud, securing cloud infrastructure becomes more critical. Cloud platforms are susceptible to a variety of threats, including data breaches, misconfigurations, insider threats, and sophisticated attacks such as hyper-jacking, where attackers exploit vulnerabilities in virtual machines. If attackers compromise the cloud infrastructure, they can gain unauthorized access to critical ICPS data or manipulate industrial processes remotely, posing significant risks to both cybersecurity and physical safety.

The increased reliance of 5G networks on software and virtualization also makes them more susceptible to MitM attacks. In a cloud-enabled ICPS, attackers could intercept communications between cloud services and industrial devices, alter data, or inject malicious commands, potentially leading to disruptions in critical operations or creating safety hazards in physical processes. Given these risks, there is an urgent need for research and solutions to protect 5G-enabled ICPS environments. One key strategy is network segmentation [138], which can isolate critical industrial systems from less secure or public networks. Micro-segmentation techniques further enhance this by containing potential attacks and preventing them from spreading across the ICPS. In addition, zero-trust security models [139] are essential to ensure that no device, user, or application is trusted by default, even if it is within the network perimeter. Continuous verification, strict access controls, and multi-factor authentication can significantly reduce the risk of unauthorized access to critical systems.

AI-driven security solutions are also pivotal in securing 5G-enabled ICPS. By monitoring network traffic, detecting anomalies, and predicting potential threats in real-time, AI can help identify and mitigate cyberattacks before they cause significant damage. In this evolving landscape, AI-based tools play an increasingly crucial role in strengthening the security and resilience of industrial control systems.

2) *Research Directions Driven by AI*: Recent advances and excitement in foundation models, such as LLMs and deep learning architectures, further increase the power of prediction and automated responses in handling security concerns [140]. These models can process and understand vast amounts of data, learning intricate patterns that simpler models might miss. By integrating foundation models into security systems, organizations can achieve more accurate threat detection and more sophisticated response strategies. For example, foundation models can improve the understanding of complex attack vectors, enabling more effective defense mechanisms that

adapt to evolving threats. The work presented in [140] explores an innovative approach to the design of foundation models for the management of cloud systems, focusing particularly on meta-learning techniques.

Recent work in [141] has introduced an innovative approach that combines integrated game-theoretic models with foundation models to improve data analytics for cybersecurity. This research aims to develop descriptive and prescriptive analytics to design proactive deception mechanisms, specifically targeting the mitigation of APT attacks. By leveraging game theory, the study provides a strategic framework for understanding the interactions between attackers and defenders, while the use of foundation models enhances the adaptability and effectiveness of these strategies. The proposed approach not only identifies potential attack vectors but also prescribes optimal countermeasures that can be implemented in real-time, significantly reducing the impact of APT attacks on critical systems. This integration of advanced modeling techniques is a promising direction in the development of resilient and proactive cybersecurity solutions.

The work presented in [142] improves the practical application of LLMs in real-world robotic tasks by proposing a novel approach that integrates LLMs with pre-trained robotic skills. This integration allows the language model to generate actions in natural language that are not only contextually appropriate but also feasible within the robot's operational environment. Essentially, the robot acts as the "hands and eyes" of the language model, translating abstract/high-level instructions into concrete/executable tasks. This method effectively marries the semantic understanding of LLMs with the grounded, environment-specific capabilities of robotic systems, enabling robots to carry out complex, long-horizon tasks guided by natural language instructions. The success of this approach has been demonstrated through rigorous evaluations of real-world robotic tasks, demonstrating its potential to significantly advance the field of human-robot interaction.

The use of AI in protecting ICPS offers tremendous potential to improve security and resilience, but also introduces its own set of vulnerabilities. AI can improve threat detection, automate responses, and provide predictive analytics to prevent security incidents in ICPS. However, the very algorithms that power AI are susceptible to attacks, such as data poisoning, where attackers introduce malicious data into the training process, causing the model to make incorrect predictions or classifications. This can have serious consequences in an industrial environment, where incorrect decisions could disrupt critical operations, cause safety hazards, and even cause damage to the physical infrastructure. Manipulation attacks on AI models, particularly LLMs, present another significant risk. Adversaries can exploit LLMs by manipulating the inputs to achieve desired but incorrect outputs, potentially misleading decision-makers, or causing automated systems to behave in unintended ways. For example, LLMs could be used to help manage complex operations or provide recommendations for system adjustments, but if an attacker manipulates the model, the results could lead to operational inefficiencies or even critical system failures.

It is crucial to be aware of these AI vulnerabilities while simultaneously developing secure and resilient solutions for ICPS. AI's increasing role in ICPS security demands a comprehensive approach that not only incorporates AI-based defenses but also actively addresses the specific threats posed to AI itself. Robust defenses such as adversarial training, data integrity checks, and secure model development processes must be used to mitigate the risks of data poisoning and model manipulation. The holistic integration of AI, cyber, and physical ICPS stacks creates additional layers of complexity. The interplay between these elements introduces new challenges in the future design of secure systems. For example, AI may monitor and control cyber elements, such as network traffic, while simultaneously interfacing with physical components, such as sensors and actuators. Ensuring that each layer, i.e., AI algorithms, cybersecurity protocols, and physical controls, works harmoniously and securely together requires a multidisciplinary cyber-secure system design approach.

VII. CONCLUDING REMARKS

Recent research and development efforts in advancing network-centric industrial automation reference architectures such as Cloud-Fog Automation have prompted us to explore its broader and more futuristic applicability. In this paper, we have presented a comprehensive overview of the foundations of Cloud-Fog Automation as the new paradigm for realizing the vision of a fully autonomous ICPS, as well as highlighting the crucial role of communication, computing, and control co-design, i.e., "3C co-design". We first introduced a Cloud-Fog Automation prototype which focused on four representative industrial use cases. We then outlined the vision and design requirements of autonomous ICPS, presented research gaps and discussed how 3C co-design is poised to close the gap between the new-generation digital infrastructure and future autonomous ICPS. We subsequently examined the state-of-the-art in goal-oriented communication, virtualization-empowered computing, QoS-aware control; and identify key opportunities and future research directions to foster innovation and bridge the gap between theory and practical applications in these inter-related domains. Finally, we emphasized the critical importance of cyber-physical security as a cross-cutting concern across the communication, computing, and control domains.

There are tremendous opportunities, open research challenges and future directions in delivering an integrated and holistic 3C co-design for autonomous ICPS. It is our sincere hope that this paper serves as both a roadmap and a call-for-action for the community to steer research and development efforts in the field, with a strong emphasis on the pivotal role of 3C co-design in driving Cloud-Fog Automation as the foundational paradigm for future autonomous ICPS.

REFERENCE

- [1] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial Internet of Things: Challenges, opportunities, and directions," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4724–4734, Nov. 2018.
- [2] W. Xiang, K. Yu, F. Han, L. Fang, D. He, and Q.-L. Han, "Advanced manufacturing in industry 5.0: A survey of key enabling technologies and future trends," *IEEE Trans. Ind. Informat.*, vol. 20, no. 2, pp. 1055–1068, Feb. 2024.
- [3] Enterprise-Control System Integration—Part 1: Models and Terminology, Standard ANSI/ISA-95.00.01-2010, International Society of Automation, 2000.
- [4] J. Jin, K. Yu, J. Kua, N. Zhang, Z. Pang, and Q.-L. Han, "Cloud-fog automation: Vision, enabling technologies, and future research directions," *IEEE Trans. Ind. Informat.*, vol. 20, no. 2, pp. 1039–1054, Feb. 2024.
- [5] H. Lyu, J. Yan, J. Zhang, Z. Pang, G. Yang, and A. J. Isaksson, "Cloud-Fog automation: Heterogenous applications over new-generation infrastructure of virtualized computing and converged networks," *IEEE Ind. Electron. Mag.*, vol. 18, no. 4, pp. 2–14, Jan. 2024.
- [6] M. Aazam, S. Zeadally, and K. A. Harras, "Deploying fog computing in industrial Internet of Things and industry 4.0," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4674–4682, Oct. 2018.
- [7] A. M. Annaswamy, K. H. Johansson, and G. J. Pappas, "Control for societal-scale challenges: Road map 2030," *IEEE Control Syst.*, vol. 44, no. 3, pp. 30–32, Jun. 2024. [Online]. Available: <https://ieeecs.org/control-societal-scale-challenges-roadmap-2030>
- [8] A. Ahlén et al., "Toward wireless control in industrial process automation: A case study at a paper mill," *IEEE Control Syst. Mag.*, vol. 39, no. 5, pp. 36–57, Oct. 2019.
- [9] B. Dafflon, N. Moalla, and Y. Ouzrout, "The challenges, approaches, and used techniques of CPS for manufacturing in industry 4.0: A literature review," *Int. J. Adv. Manuf. Technol.*, vol. 113, nos. 7–8, pp. 2395–2412, Apr. 2021, doi: [10.1007/s00170-020-06572-4](https://doi.org/10.1007/s00170-020-06572-4).
- [10] Y. Liu, X. Tao, X. Li, A. W. Colombo, and S. Hu, "Artificial intelligence in smart logistics cyber-physical systems: State-of-the-arts and potential applications," *IEEE Trans. Ind. Cyber-Phys. Syst.*, vol. 1, no. 1, pp. 1–20, Mar. 2023.
- [11] B. Besselink et al., "Cyber-Physical control of road freight transport," *Proc. IEEE*, vol. 104, no. 5, pp. 1128–1141, May 2016.
- [12] P. Park, S. Coleri Ergen, C. Fischione, C. Lu, and K. H. Johansson, "Wireless network design for control systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 978–1013, 2nd Quart., 2018.
- [13] J. Cao, E. Kurniawan, A. Boonkajay, S. Sun, P. Popovski, and X. Zhu, "Goal-oriented integration of sensing, communication, computing, and control for mission-critical Internet-of-Things," *IEEE Netw.*, vol. 38, no. 6, pp. 13–20, Nov. 2024.
- [14] E. Fridman, *Introduction To Time-Delay Systems*. Cambridge, MA, USA: Cambridge Univ. Press, 2014.
- [15] L. Li, A. E. Kalør, P. Popovski, and W. Chen, "Unified timing analysis for closed-loop goal-oriented wireless communication," 2024, [arXiv:2405.16047](https://arxiv.org/abs/2405.16047).
- [16] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *Proc. IEEE*, vol. 95, no. 1, pp. 108–137, Jan. 2007.
- [17] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proc. IEEE*, vol. 95, no. 1, pp. 163–187, Jan. 2007.
- [18] E. Uysal et al., "Semantic communications in networked systems: A data significance perspective," *IEEE Netw.*, vol. 36, no. 4, pp. 233–240, Jul./Aug. 2022.
- [19] A. Nasrallah et al., "Ultra-low latency (ULL) networks: The IEEE TSN and IETF DetNet standards and related 5G ULL research," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 88–145, 1st Quart., 2018.
- [20] D. Cavalcanti, C. Cordeiro, M. Smith, and A. Regev, "WiFi TSN: Enabling deterministic wireless connectivity over 802.11," *IEEE Commun. Standards Mag.*, vol. 6, no. 4, pp. 22–29, Dec. 2022.
- [21] P. M. Rost and T. Kolding, "Performance of integrated 3GPP 5G and IEEE TSN networks," *IEEE Commun. Standards Mag.*, vol. 6, no. 2, pp. 51–56, Jun. 2022.
- [22] S. Sudhakaran, K. Montgomery, M. Kashef, D. Cavalcanti, and R. Candell, "Wireless time sensitive networking impact on an industrial collaborative robotic workcell," *IEEE Trans. Ind. Informat.*, vol. 18, no. 10, pp. 7351–7360, Oct. 2022.
- [23] T. Adame, M. Carrascosa-Zamacois, and B. Bellalta, "Time-sensitive networking in IEEE 802.11be: On the way to low-latency WiFi 7," *Sensors*, vol. 21, no. 15, pp. 4954–4974, Jul. 2021.
- [24] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana, IL, USA: Univ. of Illinois Press, 1949.

- [25] E. C. Strinati and S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications," *Comput. Netw.*, vol. 190, May 2021, Art. no. 107930.
- [26] X. Luo, H.-H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 210–219, Feb. 2022.
- [27] K. Niu et al., "A paradigm shift toward semantic communications," *IEEE Commun. Mag.*, vol. 60, no. 11, pp. 113–119, Nov. 2022.
- [28] Z. Qin, X. Tao, J. Lu, W. Tong, and G. Ye Li, "Semantic communications: Principles and challenges," 2021, *arXiv:2201.01389*.
- [29] D. Gündüz et al., "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5–41, Jan. 2023.
- [30] T. M. Getu, G. Kaddoum, and M. Bennis, "Making sense of meaning: A survey on metrics for semantic and goal-oriented communication," *IEEE Access*, vol. 11, pp. 45456–45492, 2023.
- [31] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *Proc. 46th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2012, pp. 1–6.
- [32] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.
- [33] M. A. Abd-Elmagid, N. Pappas, and H. S. Dhillon, "On the role of age of information in the Internet of Things," *IEEE Commun. Mag.*, vol. 57, no. 12, pp. 72–77, Dec. 2019.
- [34] X. Wang, C. Chen, J. He, S. Zhu, and X. Guan, "AoI-aware control and communication co-design for industrial IoT systems," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8464–8473, May 2021.
- [35] M. H. Mamduhi, J. P. Champati, J. Gross, and K. H. Johansson, "Where freshness matters in the control loop: Mixed age-of-information and event-based co-design for multi-loop networked control systems," *J. Sensor Actuator Netw.*, vol. 9, no. 3, p. 43, Sep. 2020. [Online]. Available: <https://www.mdpi.com/2224-2708/9/3/43>
- [36] M. Salimnejad, M. Kountouris, A. Ephremides, and N. Pappas, "Age of information versions: A semantic view of Markov source monitoring," 2024, *arXiv:2406.14594*.
- [37] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age and value of information: Non-linear age case," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 326–330.
- [38] X. Zheng, S. Zhou, and Z. Niu, "Urgency of information for context-aware timely status updates in remote control systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7237–7250, Nov. 2020.
- [39] P. Agheli, N. Pappas, P. Popovski, and M. Kountouris, "Effective communication: When to pull updates?," in *Proc. ICC - IEEE Int. Conf. Commun.*, Jun. 2024, pp. 183–188.
- [40] P. M. de Sant Ana, N. Marchenko, P. Popovski, and B. Soret, "Age of loop for wireless networked control systems optimization," in *Proc. IEEE 32nd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2021, pp. 1–7.
- [41] A. Li, S. Wu, S. Meng, R. Lu, S. Sun, and Q. Zhang, "Toward goal-oriented semantic communications: New metrics, framework, and open challenges," *IEEE Wireless Commun.*, vol. 31, no. 5, pp. 238–245, Oct. 2024.
- [42] A. Li, S. Wu, and S. Sun, "Goal-oriented tensor: Beyond AoI towards semantics-empowered Goal-oriented communications," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.
- [43] H. Zou, C. Zhang, S. Lasaulce, L. Saludjian, and H. V. Poor, "Goal-oriented quantization: Analysis, design, and application to resource allocation," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 42–54, Jan. 2023.
- [44] N. Pappas and M. Kountouris, "Goal-oriented communication for real-time tracking in autonomous systems," in *Proc. IEEE Int. Conf. Auto. Syst. (ICAS)*, Aug. 2021, pp. 1–5.
- [45] C. Zhang, H. Zou, S. Lasaulce, W. Saad, M. Kountouris, and M. Bennis, "Goal-oriented communications for the IoT and application to data compression," *IEEE Internet Things Mag.*, vol. 5, no. 4, pp. 58–63, Dec. 2022.
- [46] P. Agheli, N. Pappas, and M. Kountouris, "Semantic filtering and source coding in distributed wireless monitoring systems," *IEEE Trans. Commun.*, vol. 72, no. 6, pp. 3290–3304, Jun. 2024.
- [47] J. M. Gimenez-Guzman, I. Leyva-Mayorga, and P. Popovski, "Goal-oriented source coding and filtering for vehicular communications," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19804–19822, Jun. 2024.
- [48] J. Xu, T.-Y. Tung, B. Ai, W. Chen, Y. Sun, and D. Gündüz, "Deep joint source-channel coding for semantic communications," *IEEE Commun. Mag.*, vol. 61, no. 11, pp. 42–48, Nov. 2023.
- [49] W. F. Lo, N. Mital, H. Wu, and D. Gündüz, "Collaborative semantic communication for edge inference," *IEEE Wireless Commun. Lett.*, vol. 12, no. 7, pp. 1125–1129, Jul. 2023.
- [50] T.-Y. Tung and D. Gündüz, "DeepWiVe: Deep-Learning-Aided wireless video transmission," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2570–2583, Sep. 2022.
- [51] E. T. Ceran, D. Gündüz, and A. György, *Reinforcement Learning for Minimizing Age of Information Over Wireless Links*. Cambridge, U.K.: Cambridge Univ. Press, 2023, pp. 327–363.
- [52] F. Mason, F. Chiarioti, A. Zanella, and P. Popovski, "Multi-agent reinforcement learning for coordinating communication and control," *IEEE Trans. Cognit. Commun. Netw.*, vol. 10, no. 4, pp. 1566–1581, Aug. 2024.
- [53] P. Talli et al., "Pragmatic communication for remote control of finite-state Markov processes," 2024, *arXiv:2405.10672*.
- [54] F. Pezone, S. Barbarossa, and P. Di Lorenzo, "Goal-oriented communication for edge learning based on the information bottleneck," in *Proc. ICASSP - IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2022, pp. 8832–8836.
- [55] T.-Y. Tung, S. Kobus, J. P. Roig, and D. Gündüz, "Effective communications: A joint learning and communication framework for multi-agent reinforcement learning over noisy channels," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2590–2603, Aug. 2021.
- [56] P. Talli, F. Mason, F. Chiarioti, and A. Zanella, "Push- and pull-based effective communication in cyber-physical systems," 2024, *arXiv:2401.10921*.
- [57] J. N. Foerster, Y. M. Assael, N. de Freitas, and S. Whiteson, "Learning to communicate with deep multi-agent reinforcement learning," *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, pp. 2145–2153, Dec. 2016.
- [58] A. Das et al., "TarMAC: Targeted multi-agent communication," in *Proc. Int. Conf. Mach. Learn.*, May 2019, pp. 1538–1546.
- [59] M. Chafii, S. Naoumi, R. Alami, E. Almazrouei, M. Bennis, and M. Debbah, "Emergent communication in multi-agent reinforcement learning for future wireless networks," *IEEE Internet Things Mag.*, vol. 6, no. 4, pp. 18–24, Dec. 2023.
- [60] S. Mostafa, M. P. Mota, A. Valcarce, and M. Bennis, "Emergent communication protocol learning for task offloading in industrial Internet of Things," in *Proc. GLOBECOM - IEEE Global Commun. Conf.*, Dec. 2023, pp. 7055–7060.
- [61] M. Merluzzi, M. C. Filippou, L. Gomes Baltar, M. D. Mueck, and E. Calvanese Strinati, "6G goal-oriented communications: How to coexist with legacy systems?," *Telecom*, vol. 5, no. 1, pp. 65–97, Jan. 2024.
- [62] P. Gaj, M. Skrzewski, J. Stój, and J. Flak, "Virtualization as a way to distribute PC-based functionalities," *IEEE Trans. Ind. Informat.*, vol. 11, no. 3, pp. 763–770, Jun. 2015.
- [63] M. Pearce, S. Zeadally, and R. Hunt, "Virtualization: Issues, security threats, and solutions," *ACM Comput. Surveys*, vol. 45, no. 2, pp. 1–39, Feb. 2013.
- [64] J. Li, F. R. Yu, G. Deng, C. Luo, Z. Ming, and Q. Yan, "Industrial internet: A survey on the enabling technologies, applications, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1504–1526, 3rd Quart., 2017.
- [65] P. Pop et al., "The FORA fog computing platform for industrial IoT," *Inf. Syst.*, vol. 98, pp. 101727–101753, May 2021, doi: [10.1016/j.is.2021.101727](https://doi.org/10.1016/j.is.2021.101727).
- [66] N. Mahmud, K. Sandström, and A. Vulgarakis, "Evaluating industrial applicability of virtualization on a distributed multicore platform," in *Proc. IEEE Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2014, pp. 1–8.
- [67] M. Cinque, D. Cotroneo, L. De Simone, and S. Rosiello, "Virtualizing mixed-criticality systems: A survey on industrial trends and issues," *Future Gener. Comput. Syst.*, vol. 129, pp. 315–330, Apr. 2022, doi: [10.1016/j.future.2021.12.002](https://doi.org/10.1016/j.future.2021.12.002).
- [68] M. Gundall, D. Reti, and H. D. Schotten, "Application of virtualization technologies in novel industrial automation: Catalyst or show-stopper?," in *Proc. IEEE 18th Int. Conf. Ind. Informat. (INDIN)*, vol. 1, Jul. 2020, pp. 283–290. [Online]. Available: <https://ieeexplore.ieee.org/document/9442123/>
- [69] J. Ruh and W. Steiner, "The need for deterministic virtualization in the industrial Internet of Things," in *Proc. Workshop Fog Comput. IoT (IoT-Fog)*, Apr. 2019, pp. 26–30.
- [70] A. Moga, T. Sivanthi, and C. Franke, "OS-level virtualization for industrial automation systems: Are we there yet?," in *Proc. 31st Annu. ACM Symp. Appl. Comput.*, Apr. 2016, pp. 1838–1843.

- [71] L. Yin, J. Luo, and H. Luo, "Tasks scheduling and resource allocation in fog computing based on containers for smart manufacturing," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4712–4721, Oct. 2018.
- [72] M. Barletta, M. Cinque, L. De Simone, and R. D. Corte, "Criticality-aware monitoring and orchestration for containerized industry 4.0 environments," *ACM Trans. Embedded Comput. Syst.*, vol. 23, no. 1, pp. 1–28, Jan. 2024. [Online]. Available: <https://dl.acm.org/doi/10.1145/3604567>
- [73] T. Goldschmidt, S. Hauck-Stattelmann, S. Malakuti, and S. Grüner, "Container-based architecture for flexible industrial control applications," *J. Syst. Archit.*, vol. 84, pp. 28–36, Mar. 2018.
- [74] M. Sollfrank, F. Loch, S. Denteneer, and B. Vogel-Heuser, "Evaluating Docker for lightweight virtualization of distributed and time-sensitive applications in industrial automation," *IEEE Trans. Ind. Informat.*, vol. 17, no. 5, pp. 3566–3576, May 2021.
- [75] R. Morabito, I. Farris, A. Iera, and T. Taleb, "Evaluating performance of containerized IoT services for clustered devices at the network edge," *IEEE Internet Things J.*, vol. 4, no. 4, pp. 1019–1030, Aug. 2017.
- [76] M. Becattini et al., "Empirical application insights on industrial data and service aspects of digital twin networks," in *Proc. IEEE Int. Medit. Conf. Commun. Netw. (MeditCom)*, Jul. 2024, pp. 323–328.
- [77] J. Thönes, "Microservices," *IEEE Softw.*, vol. 32, no. 1, p. 116, Jan. 2015.
- [78] N. Dragoni et al., "Microservices: Yesterday, today, and tomorrow," in *Proc. Present Superior Softw. Eng.*, Jan. 2017, pp. 195–216.
- [79] W. Dai, Y. Zhang, L. Kong, J. H. Christensen, and D. Huang, "Design of industrial edge applications based on IEC 61499 microservices and containers," *IEEE Trans. Ind. Informat.*, vol. 19, no. 7, pp. 7925–7935, Jul. 2023.
- [80] K. Thramboulidis, D. C. Vachtsevanou, and A. Solanos, "Cyber-physical microservices: An IoT-based framework for manufacturing systems," *Proc. IEEE Ind. Cyber-Phys. Syst. (ICPS)*, vol. 50, pp. 232–239, May 2018.
- [81] A. Homain, M. de Sousa, A. Zoitl, and M. Wollschlaeger, "Service granularity in industrial automation and control systems," in *Proc. 25th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, vol. 1, Sep. 2020, pp. 132–139.
- [82] H. Koziolok, A. Burger, P. P. Abdulla, J. Rückert, S. Sonar, and P. Rodriguez, "Dynamic updates of virtual PLCs deployed as Kubernetes microservices," in *Proc. Eur. Conf. Softw. Archit.*, Jan. 2021, pp. 3–19.
- [83] T. Lyu, U. Dwi Atmojo, and V. Vyatkin, "Towards cloud-based virtual commissioning of distributed automation applications with IEC 61499 and containerization technology," in *Proc. 47th Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2021, pp. 1–7.
- [84] B. Han, V. Gopalakrishnan, L. Ji, and S. Lee, "Network function virtualization: Challenges and opportunities for innovations," *IEEE Commun. Mag.*, vol. 53, no. 2, pp. 90–97, Feb. 2015.
- [85] J. Li, X. Qi, J. Li, Z. Su, Y. Su, and L. Liu, "Fault diagnosis in the network function virtualization: A survey, taxonomy, and future directions," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19121–19142, Jun. 2024.
- [86] L. Leonardi, M. Ashjaei, H. Fotouhi, and L. L. Bello, "A proposal towards software-defined management of heterogeneous virtualized industrial networks," in *Proc. IEEE 17th Int. Conf. Ind. Informat. (INDIN)*, vol. 1, Jul. 2019, pp. 1741–1746.
- [87] I. Farris, T. Taleb, Y. Khettab, and J. Song, "A survey on emerging SDN and NFV security mechanisms for IoT systems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 812–837, 1st Quart., 2019.
- [88] F. Bannour, S. Souihi, and A. Mellouk, "Distributed SDN control: Survey, taxonomy, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 333–354, 1st Quart., 2017.
- [89] H. R. Chi, C. K. Wu, N.-F. Huang, K.-F. Tsang, and A. Radwan, "A survey of network automation for industrial Internet-of-Things toward industry 5.0," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 2065–2077, Feb. 2023.
- [90] L. Tang, W. Wen, J. Li, D. Fang, L. Li, and Q. Chen, "Digital-Twin-Assisted VNF mapping and scheduling in SDN/NFV-enabled industrial IoT," *IEEE Internet Things J.*, vol. 11, no. 10, pp. 18516–18533, May 2024. [Online]. Available: <https://ieeexplore.ieee.org/document/10430364/>
- [91] J. Liang and F. Tian, "An online algorithm for virtualized network function placement in mobile edge industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 19, no. 3, pp. 2496–2507, Mar. 2023.
- [92] G. Marchetto, R. Sisto, F. Valenza, J. Yusupov, and A. Ksentini, "A formal approach to verify connectivity and optimize VNF placement in industrial networks," *IEEE Trans. Ind. Informat.*, vol. 17, no. 2, pp. 1515–1525, Feb. 2021.
- [93] P. Rahimi, C. Chrysostomou, H. Pervaiz, V. Vassiliou, and Q. Ni, "Joint radio resource allocation and beamforming optimization for industrial Internet of Things in software-defined networking-based virtual fog-radio access network 5G-and-beyond wireless environments," *IEEE Trans. Ind. Informat.*, vol. 18, no. 6, pp. 4198–4209, Jun. 2022.
- [94] P.-V. Mekikis et al., "NFV-enabled experimental platform for 5G tactile Internet support in industrial environments," *IEEE Trans. Ind. Informat.*, 2020, vol. 16, no. 3, pp. 1895–1903, Mar. 2020.
- [95] H. Huang, W. Miao, G. Min, J. Tian, and A. Alamri, "NFV and blockchain enabled 5G for ultra-reliable and low-latency communications in industry: Architecture and performance evaluation," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5595–5604, Aug. 2021.
- [96] M. Du and K. Wang, "An SDN-enabled pseudo-honeypot strategy for distributed denial of service attacks in Industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 648–657, Jan. 2020.
- [97] L. Ji, S. He, W. Wu, C. Gu, J. Bi, and Z. Shi, "Dynamic network slicing orchestration for remote adaptation and configuration in industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 18, no. 6, pp. 4297–4307, Jun. 2022.
- [98] T. Mai, H. Yao, N. Zhang, W. He, D. Guo, and M. Guizani, "Transfer reinforcement learning aided distributed network slicing optimization in Industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 18, no. 6, pp. 4308–4316, Jun. 2022.
- [99] H. Wang, Y. Wu, G. Min, and W. Miao, "A graph neural network-based digital twin for network slicing management," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1367–1376, Feb. 2022.
- [100] A. Fischer, J. F. Botero, M. T. Beck, H. De Meer, and X. Hesselbach, "Virtual network embedding: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1888–1906, 4th Quart., 2013.
- [101] X. Xiao, "DVNE-DRL: Dynamic virtual network embedding algorithm based on deep reinforcement learning," *Sci. Rep.*, vol. 13, no. 1, p. 19789, Nov. 2023.
- [102] H. Cao, S. Wu, G. S. Aujla, Q. Wang, L. Yang, and H. Zhu, "Dynamic embedding and quality of service-driven adjustment for cloud networks," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1406–1416, Feb. 2020.
- [103] M. Li, C. Chen, C. Hua, and X. Guan, "Intelligent latency-aware virtual network embedding for industrial wireless networks," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7484–7496, Oct. 2019.
- [104] M. Li, C. Hua, C. Chen, and X. Guan, "Application-driven virtual network embedding for industrial wireless sensor networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.
- [105] K. Makhijani and L. Dong. (2022). *Virtualization of PLC in Industrial Networks-Problem Statement*. Internet Engineering Task Force (IETF). [Online]. Available: <https://datatracker.ietf.org/doc/draft-km-iotops-iiot-fwk/02/>
- [106] M. Azarmipour, H. Elfaham, C. Gries, and U. Epple, "PLC 4.0: A control system for industry 4.0," in *Proc. 45th Annu. Conf. IEEE Ind. Electron. Soc.*, vol. 1, Oct. 2019, pp. 5513–5518.
- [107] M. Gaffurini, P. Bellagente, A. Depari, A. Flammini, E. Sisinni, and P. Ferrari, "Virtual PLC in industrial edge platform: Performance evaluation of supervision and control communication," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–10, 2024.
- [108] C. Scordino et al., "Real-time virtualization for industrial automation," in *Proc. 25th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, vol. 1, Sep. 2020, pp. 353–360.
- [109] Q. Yu, H. Wan, X. Zhao, Y. Gao, and M. Gu, "Online scheduling for dynamic VM migration in multicast time-sensitive networks," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 3778–3788, Jun. 2020.
- [110] L. Orciari, D. Raggini, and A. Tilli, "Taming edge computing for hard real-time advanced control of mechatronic systems," *IEEE Trans. Ind. Informat.*, vol. 20, no. 8, pp. 9898–9906, Aug. 2024.
- [111] A. Cilaro, M. Cinque, L. De Simone, and N. Mazzocca, "Virtualization over multiprocessor systems-on-chip: An enabling paradigm for the industrial Internet of Things," *Computer*, vol. 55, no. 10, pp. 35–47, Oct. 2022.
- [112] D. Wu, Z. Li, J. Wang, Y. Zheng, M. Li, and Q. Huang, "Vision and challenges for knowledge centric networking," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 117–123, Aug. 2019.
- [113] M. Chiang and T. Zhang, "Fog and IoT: An overview of research opportunities," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 854–864, Dec. 2016.

- [114] A. Botta, W. de Donato, V. Persico, and A. Pescapé, "Integration of cloud computing and Internet of Things: A survey," *Future Gener. Comput. Syst.*, vol. 56, pp. 684–700, Mar. 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167739X15003015>
- [115] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, Jan. 2017.
- [116] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0005109899002149>
- [117] B. Cao, Y. Fu, Z. Sun, X. Liu, H. He, and Z. Lv, "A resource allocation strategy in fog-cloud computing towards the Internet of Things in the 5G era," in *Proc. IEEE 26th Int. Workshop Comput. Aided Model. Design Commun. Links Netw. (CAMAD)*, Oct. 2021, pp. 1–6.
- [118] Y. Liu, P. Zeng, J. Cui, C. Xia, and Y. Sun, "A self-triggered approach for co-design of MPC and computing resource allocation," *IEEE Internet Things J.*, vol. 11, no. 14, pp. 25024–25032, Jul. 2024.
- [119] T. Zhu, J. Mao, L. Han, C. Zhang, and J. Yang, "Real-time dynamic obstacle avoidance for robot manipulators based on cascaded nonlinear MPC with artificial potential field," *IEEE Trans. Ind. Electron.*, vol. 71, no. 7, pp. 7424–7434, Jul. 2024.
- [120] Z. Wang, R. Liu, Q. Liu, L. Han, Y. Wu, and J. S. Thompson, "QoS-oriented sensing–communication–control co-design for UAV-enabled positioning," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 1, pp. 497–511, Mar. 2023.
- [121] H. Zhao, X. Dai, P. Zhou, and T. Yang, "Distributed robust event-triggered control strategy for multiple high-speed trains with communication delays and input constraints," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 3, pp. 1453–1464, Sep. 2020.
- [122] X. Wang, J. Sun, G. Wang, F. Allgöwer, and J. Chen, "Data-driven control of distributed event-triggered network systems," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 2, pp. 351–364, Feb. 2023.
- [123] H. Lyu, Z. Pang, A. Bengtsson, S. Nilsson, A. J. Isaksson, and G. Yang, "Latency-aware control for wireless cloud-fog automation: Framework and case study," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 5400–5410, 2024.
- [124] S. Zhou, Z. Jiang, N. Pappas, A. Ephremides, and L. A. DaSilva, "Guest editorial special issue on age of information and data semantics for sensing, communication, and control co-design in IoT," *IEEE Internet Things J.*, vol. 8, no. 19, pp. 14431–14434, Oct. 2021.
- [125] I. Martínez, A. S. Hafid, and A. Jarray, "Design, resource management, and evaluation of fog computing systems: A survey," *IEEE Internet Things J.*, vol. 8, no. 4, pp. 2494–2516, Feb. 2021.
- [126] Y. Wan, C. Long, R. Deng, G. Wen, and X. Yu, "Adaptive event-triggered strategy for economic dispatch in uncertain communication networks," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 4, pp. 1881–1891, Dec. 2021.
- [127] Y. Cai, J. Llorca, A. M. Tulino, and A. F. Molisch, "Ultra-reliable distributed cloud network control with end-to-end latency constraints," *IEEE/ACM Trans. Netw.*, vol. 30, no. 6, pp. 2505–2520, Dec. 2022.
- [128] X. Hu, G. Xing, and J. Y.-T. Leung, "Exploring the interplay between computation and communication in distributed real-time scheduling," *IEEE Trans. Comput.*, vol. 60, no. 12, pp. 1759–1771, Dec. 2011.
- [129] C. Tang, G. Yan, H. Wu, and C. Zhu, "Computation offloading and resource allocation in failure-aware vehicular edge computing," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 1877–1888, Feb. 2024.
- [130] X. Li, X. Liang, R. Lu, X. Shen, X. Lin, and H. Zhu, "Securing smart grid: Cyber attacks, countermeasures, and challenges," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 38–45, Aug. 2012.
- [131] M. Ali, S. U. Khan, and A. V. Vasilakos, "Security in cloud computing: Opportunities and challenges," *Inf. Sci.*, vol. 305, pp. 357–383, Jun. 2015.
- [132] Q. Zhu and L. Bushnell, "Networked cyber-physical systems: Interdependence, resilience and information exchange," in *Proc. 51st Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2013, pp. 763–769.
- [133] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints*. Cambridge, MA, USA: Cambridge Univ. Press, 2013.
- [134] J. Moon and T. Başar, "Control over lossy networks: A dynamic game approach," *Proc. Amer. Control Conf. (ACC)*, pp. 5367–5372, Jun. 2014.
- [135] X.-M. Zhang, Q.-L. Han, and X. Yu, "Survey on recent advances in networked control systems," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1740–1752, Oct. 2016.
- [136] Q. Zhu and T. Başar, "Disentangling resilience from robustness: Contextual dualism, interactionism, and game-theoretic paradigms," *IEEE Control Syst.*, vol. 44, no. 3, pp. 95–103, Jun. 2024.
- [137] Z. Xu and Q. Zhu, "Secure and resilient control design for cloud enabled networked control systems," in *Proc. 1st ACM Workshop Cyber-Physical Syst.-Secur. Privacy*, Oct. 2015, pp. 31–42.
- [138] M. Chahbar, G. Diaz, A. Dandoush, C. Christophe, and K. Ghomid, "A comprehensive survey on the E2E 5G network slicing model," *IEEE Trans. Netw. Service Manag.*, vol. 4537, no. 1, pp. 49–62, Mar. 2020.
- [139] V. Stafford, "Zero trust architecture," *NIST Special Publication*, vol. 800, p. 207, Aug. 2020.
- [140] H. Qiu, W. Mao, C. W. H. Franke, Z. T. Kalbarczyk, T. Basar, and R. K. Iyer, "On the promise and challenges of foundation models for learning-based cloud systems management," in *Proc. Workshop Mach. Learn. Syst. NeurIPS*, 2023, pp. 1–21.
- [141] T. Li and Q. Zhu, "Symbiotic game and foundation models for cyber deception operations in strategic cyber warfare," 2024, [arXiv:2403.10570](https://arxiv.org/abs/2403.10570).
- [142] M. J. Ahn et al., "Do as I can, not as I say: Grounding language in robotic affordances," in *Proc. Conf. Robot Learn. (CoRL)*, Jan. 2022, pp. 287–318.



Jiong Jin (Member, IEEE) received the B.E. degree with First Class Honours in computer engineering from Nanyang Technological University, Singapore, in 2006, and the Ph.D. degree in electrical and electronic engineering from the University of Melbourne, Australia, in 2011. He is currently a Full Professor with the School of Engineering, Swinburne University of Technology, Melbourne, Australia. His research interests include network design and optimization, edge computing and intelligence, robotics and automation, and cyber-physical systems and Internet of Things as well as their applications in smart manufacturing, smart transportation, and smart cities. He was recognized as an Honorable Mention in the AI 2000 Most Influential Scholars List in IoT (2021 and 2022). He is currently an Associate Editor of IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS and IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING.



Zhibo Pang (Senior Member, IEEE) received the M.B.A. degree in innovation and growth from the University of Turku in 2012 and the Ph.D. degree in electronic and computer systems from the KTH Royal Institute of Technology in 2013. He is a Senior Principal Scientist with ABB Corporate Research Sweden; and an Adjunct Professor with the KTH Royal Institute of Technology. He was an Adjunct Professor with The University of Sydney. He has many productized research results and 23 granted patents in U.S., Europe, and Japan. He received the Inventor of the Year Award by ABB Corporate Research Sweden, in 2016, 2018, and 2021, respectively. He works on embodied intelligence, robotics, control, computing, communication, and electronics for Industry 4.0 and Healthcare 4.0. He is a member of the IEEE IES Industry Activities Committee, a Steering Committee Member of the IEEE IoT Technical Community, the Chair of the IEEE Technical Committee on Cloud and Wireless Systems for Industrial Applications, and the Co-Chair of the IEEE TC on Industrial Informatics. He is an Associate Editor of IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, IEEE TRANSACTIONS ON SUSTAINABLE COMPUTING, IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN INDUSTRIAL ELECTRONICS, and *IEEE Internet of Things Magazine*.



Jonathan Kua (Member, IEEE) received the B.Eng. degree (Hons.) in telecommunications and network engineering and the Ph.D. degree in telecommunications engineering from Swinburne University of Technology, Australia, in 2014 and 2019, respectively. He is currently a Senior Lecturer in the Internet of Things with the School of Information Technology, Deakin University, Australia. His research interests include low-latency data transport protocols, adaptive streaming, content delivery, the Internet of Things, and industrial cyber-physical systems, with an overarching focus on improving their network performance and quality of service. He was a recipient of the Netflix Ph.D. Scholarship Award from 2015 to 2019 and the Second Runner-Up of the DASH-IF Best Ph.D. Dissertation Award on Algorithms and Protocols for Adaptive Content Delivery Over the Internet in 2019.



Quanyan Zhu (Senior Member, IEEE) is an Associate Professor with the Department of Electrical and Computer Engineering, New York University (NYU). He is an affiliated Faculty Member with the Center for Urban Science and Progress (CUSP) and the Center for Cyber Security (CCS), NYU. His research interests include cyber and physical systems, multi-agent systems, and cybersecurity and resilience. He currently serves as the Technical Committee Chair on security and privacy for the IEEE Control Systems Society. He serves as an

Associate Editor for IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS and IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING.



Karl H. Johansson (Fellow, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree in automatic control from Lund University. He is a Swedish Research Council Distinguished Professor in electrical engineering and computer science with the KTH Royal Institute of Technology, Sweden; and the Founding Director of Digital Futures. He has held visiting positions with UC Berkeley, Caltech, NTU, and other prestigious institutions. His research interests focus on networked control systems and cyber-physical systems with applications in transportation, energy, and automation networks. He is a fellow of the Royal Swedish Academy of Engineering Sciences. For his scientific contributions, he has received several best paper awards and various other distinctions from IEEE, IFAC, and other organizations. He has been awarded as

a Distinguished Professor by Swedish Research Council, a Wallenberg Scholar by the Knut and Alice Wallenberg Foundation, and a Future Research Leader by Swedish Foundation for Strategic Research. He has also received the Triennial IFAC Young Author Prize, the IEEE CSS Distinguished Lecturer, the IFAC Outstanding Service Award, and the IEEE CSS Hendrik W. Bode Lecture Prize. His extensive service to the academic community, includes being the President of European Control Association; the IEEE CSS Vice President Diversity, Outreach and Development; and a member of IEEE CSS Board of Governors and IFAC Council. He has served on the editorial boards of *Automatica*, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS, and many other journals. He has been a member of Swedish Scientific Council for Natural Sciences and Engineering Sciences.



Nikolaj Marchenko (Member, IEEE) received the Diploma degree in computer engineering from RWTH Aachen University, Germany, in 2007, and the Ph.D. degree in information technology from Klagenfurt University, Austria, in 2013. He is a Researcher Engineer with Robert Bosch GmbH, Stuttgart, Germany, where he works on the topics of the future wireless networked systems in industrial automation and automotive domains, with particular focus on joint optimization and co-design.



Dave Cavalcanti (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering from the University of Cincinnati in 2006. He is a Principal Engineer with Intel Corporation with extensive experience in distributed networked systems, wireless connectivity, industry standards, and ecosystems. He joined Intel in 2015. He is also the President of the Avnu Alliance, an industry forum driving standards and certification programs to enable deterministic performance and time sensitive networking (TSN) across ethernet/Wi-Fi/5G connectivity. He has contributed to several generations of the IEEE 802.11/Wi-Fi standards. He has published over 50 peer-reviewed papers and holds 99 granted patents. He holds several leadership positions in IEEE conferences and publications. He was a recipient of the Best Paper Awards at IEEE WFC 2022, WFC 2021, and IEEE INDIN 2021; and The Best Demo Award at IEEE INFOCOM 2018.