SoK: Applying Blockchain Technology in Industrial Internet of Things

Gang Wang
Email: email.gang.wang@gmail.com

Abstract—The proliferation of the Internet of Things (IoT) technology has made ubiquitous computing a reality by broadening Internet connectivity across diverse application domains, thus bridging billions of devices and human beings as well for information collection, data processing, and decision-making. In recent years, IoT technology and its applications in various industrial sectors have grown exponentially. Most existing industrial IoT (IIoT) implementations, however, are still relying on a centralized architecture, which is vulnerable to the single point of failure attack and requires a massive amount of computation at the central entity. The emerging blockchain technology is currently undergoing rapid development and has the full potential to revolutionize the IIoT platforms and applications. As a distributed and decentralized tamper-resistant ledger, blockchain maintains the consistency of data records at different locations and holds the potential to address the issues in traditional IIoT networks, such as heterogeneity, interoperability, and security. Integrating the blockchain technology into IIoT platforms requires to address several critical challenges that are inherent in IIoT and blockchain themselves, such as standardization, scalability, and interoperability. This paper provides a comprehensive review on the recent advances in architecture design and technology development towards tackling these challenges. We further provide several representative industrial use cases that can benefit from the integration of blockchain technology, and discuss the recent research trends and open issues in blockchain-enabled IIoT platforms.

Index Terms—Blockchain, Industrial IoT, SoK

I. INTRODUCTION

As an emerging technology, the Internet of Things (IoT) is becoming a substantial research and development area [1]. It aims at bridging billions and trillions of devices and human beings as well for fast, reliable and secure information collection, data processing, and decision-making. Gartner’s latest study on IoT industry analysis predicts that the Internet-connected things have grown to 20 billion by 2020 and will be 75 billion by 2025 [2] [3]. With the concept of Industry 4.0 [4], IoT technology is also penetrating into various industrial sectors, interconnecting sensors, actuators, controllers, along with many things together. This creates a new field – Industrial IoT (IIoT). IIoT focuses on the use of IoT to integrate emerging technologies into traditional industrial processes, such as smart sensors, robotics, machine-to-machine (M2M) interaction, big data analysis, and artificial intelligence [5] [6]. The increasing use of IIoT is expected to create new smart industrial enterprises and build the next generation smart systems [7].

IIoT platforms provide many prominent advantages to industry sectors, such as connectivity, intelligent big data analysis, edge and cloud computing, and application development. It has a significant impact on the existing industry models in many fields, including manufacturing, power system, transportation, agriculture, supply chain, and the food industry. However, as the number of IIoT devices continually increases, these devices produce huge amounts of data, which will result in much higher operational and management costs. Also, the numerous connections among devices increase some issues among device manufactures and smart factories (e.g., raising significant challenges to the interoperability, privacy, security, and fault-tolerance of IIoT).

A blockchain-based decentralized system typically stores data and information in an immutable manner, which does not require some centralized entities to control and manage these information. Emerging blockchain technological advances and applications have earned tremendous attention from both industrial and academic domains, promising to change all aspects of the digital business of the industry and solve many inherent IIoT challenges, such as interoperability and heterogeneity [8]. From a high-level perspective, blockchain is a kind of Decentralized Ledger Technology (DLT) that heavily relies on cryptographic primitives and a well-organized chain structure to securely host applications, store data, and exchange information, in an immutable and verifiable manner [9]. It is believed that the blockchain technology will have a profound impact on existing IIoT infrastructures.

As with cloud computing, big data analytics, and other new generations of information technologies, blockchain is not just one single technology; instead, it relies on many existing technologies, as well as their innovative compositions and creations, to discover and realize new capabilities. The distributed and decentralized feature, for instance, allows nodes to achieve self-management, while the centralized infrastructures in current IIoT scenarios are less efficient and are subject to various attacks, e.g., the single point of failure and DDoS attacks. The trends in decentralization can reduce IIoT’s operational and management costs. However, as of now, blockchain has several trial applications only in specific areas, such as financial services, supply chain managements [10], digital asset transactions [11], Internet of Things [12] [13] [14], and smart manufacturing [15]. Few use cases directly target practical industrial applications. Driven by the concept of Industry 4.0 [16], the blockchain-enabled IIoT platform will play a key role in reshaping various industrial applications, including manufacturing, transportation, energy management, logistics, retail, supply chain, and healthcare, to name a few.

Blockchain will serve as a driven force to enrich industrial applications. Inspired by its potential opportunities, this paper provides a systematical and comprehensive review on the integration of blockchain into industrial IoT applications. This paper covers most critical techniques for both industrial IoT and blockchain, in terms of detailed technologies, operational schemes, major challenges, and potential issues. We also provide a practical integration architecture to integrate blockchain into existing IIoT platforms, as a blockchain-enabled IIoT platform. We further provide several representative industrial use cases that can benefit from the integration of blockchain technology, and discuss the recent research trends and open issues in blockchain-enabled IIoT platforms.

The rest of this paper is organized as follows. Section II and Section III describe the features of industrial IoT and blockchain, respectively. Section IV discusses the integration of blockchain and IIoT, including the motivation, some potential integrated architectures. Section V discusses the challenges and solutions for this integration. Section VI studies several representative industrial use cases. Section VII provides some discussion on Blockchain-as-a-Service platforms. Section VIII shares our vision on the potential research trends, and Section IX concludes this paper.

G. Wang was with the University of Connecticut, Storrs, CT 06269 USA. E-mail: (gang.wang@uconn.edu)
II. FEATURES AND CHALLENGES OF INDUSTRIAL INTERNET OF THINGS (IIoT)

The modern industry is undergoing a paradigm shift from conventional computer-aided schemes to smart factories, which is empowered by recent technological advances, e.g., IoT, Artificial Intelligence (AI), and Big Data Analytics. During this shift, the IoT technology plays a key role in bridging the gap between the operational technologies (OT) deployed in existing physical industrial settings and information technologies (IT) that form the cyberspace of smart factories. This section presents the key features of IoT, and gives a summary of the major challenges that IIoT systems are facing to.

A. Industry 4.0

When we discuss IIoT, it is worth mentioning the concept of Industry 4.0 (the Fourth Industrial Revolution). Industry 4.0 is originally defined in Germany and it has gained global recognizability, which uses Internet technologies to improve production efficiency by means of smart services in smart factories. The concept of Industry 4.0 arises when the IoT paradigm is merged with the Cyber-Physical System (CPS) idea [17]. One of the preliminary goals of Industry 4.0 is generating, transmitting, and analyzing data without any interruption from a third party, as well as incorporating advanced technologies into industry sectors [4]. While the formal definition of Industry 4.0 is still in the wild, technically, IoT, IIoT, and Industry 4.0 are closely related concepts but cannot be interchangeably used. For example, IoT is often considered as a sort of web for the machines, highlighting the aim of allowing things to exchange data; while IIoT is about connecting all industrial assets, including machines and control systems that may be associated with different information systems and business processes [18].

Fig. 1(a) shows some core components of Industry 4.0, and Fig. 1(b) shows main features required for the Industry 4.0. We briefly discuss several key components consisting of a typical Industry 4.0, for example, CPS, IoT, and IoS.

a) Cyber-Physical Systems (CPS): A typical CPS extends real-world, physical objects by interconnecting them together and providing their digital descriptions. The information, stored in models and data objects, can be updated in real-time, which represents a second identity of an object itself and constitutes a sort of “digital twin” [19]. Together with the dynamic nature of these digital twins, various innovative services that were not possible in the past can be implemented across the whole product life cycle, e.g., from inception to disposal of manufactured products [18].

b) Internet of Things (IoT): IoT connects “things” (e.g., objects and machines) into the Internet, and it conceptually has some similarities with CPS. The major difference between CPS and IoT is that all IoT devices are CPS devices; however, not all CPS devices are necessarily connected to the Internet, and thus are not necessarily IoT devices. Industrial IoT is an extended version of IoT with special features and requirements, specially designed for the applications of Industry 4.0 [17].

c) Internet of Services (IoS): The main idea of IoS is that by dividing the whole section (e.g., manufacturing) into smaller components, and each then turns a simple product/component into services [20]. A product automatically evaluates the user’s expectation and transforms itself as a service that generates value (e.g., increasing the customers’ satisfaction).

Besides the above necessary core components, Industry 4.0 involves other technologies and services, for example, Big Data Analytics (BDA), Global Positioning System (GPS), Machine-to-Machine (M2M) communication, Augmented Reality (AR), and Virtual Reality (VR), as well as Artificial Intelligence and robotics [21]. Each component has its unique function in Industry 4.0 paradigm. For instance, the blockchain and decentralized ledgers can provide an immutable, secure, and decentralized transaction facility on data records.

As stated previously, IoT, IIoT, and Industry 4.0 cannot be interchangeably used. Depending on the intended goals and end-users, what is typically addressed in IoT could be better named as consumer IoT (as opposed to IIoT) [22]. In general, its communication model can be classified as machine-to-user and in the form of client-server interactions. While IIoT is about connecting all industrial assets, including machines and control systems, with associated information systems and business processes. And, the underlying communication model of IIoT is machine-oriented, and can range across a large variety of different market sectors and activities. Roughly speaking, IIoT is a subset of IoT which is specific to industrial applications, and both IoT and IIoT have a close relation with Industry 4.0 [18]. With this kind of relationship, we will discuss some unique features of industrial IoT.

B. Features of Industrial IoT

Industrial IoT network prevails with the ability to interconnect numerous devices, possessing various sensing data, with less human interventions in industry [23]. Sensing and actuating devices together form a heterogeneous industrial IoT network for various industrial applications, including manufacturing, supply chain, food industry, smart grid, healthcare, and internet of vehicles.

For different applications, the use of end-devices, communication technologies, and networking topologies may differ, especially on mobility and heterogeneity, because they must comply with regulations and demands of various applications. The topologies of industrial applications can vary with different scenarios. The classic applications with stable and mobile topologies are the industrial manufacturing production line and the vehicular ad hoc networks (VANETs) for smart transportation, respectively [24]. For example, most equipment and devices in an industrial production line are stable and are comprised of stable network topology, while vehicles in transportation move rapidly and lead to a time-varying topology. Typically, end devices with these mobilities make the network connectivity unpredictable and the entities’ management more challenging [25] [26]. The second key characteristic of IIoT is heterogeneity. IIoT devices are typically heterogeneous (e.g., with different hardware platforms and capabilities). For example, some IoT devices, such as sensors, have limited resources for processing, communication, and storage.

Although heterogeneity exists among end devices and various protocols deployed, IIoT networks still have some common features.

a) Enormous number of devices: The number of IIoT devices will continuously increase. The total number of connected devices in IoT applications is expected to increase up to 75 billion by 2025, and industrial IoT makes up to more than 17% of the number of IoT devices worldwide [27]. IIoT faces not only the issue of a large number of devices but also a growing demand for their capacities, as numerous end devices are required to sense and collect these mass data.

b) Decentralization: Decentralization is essential. Given the fact that a huge number of IIoT devices exist, such as in a supply chain application, decentralization is necessary for simultaneously processing the considerable amount of data in these devices [28]. If IIoT collects, processes, and stores these data in a decentralized manner, it can potentially mitigate the issue of a centralized bottleneck. Also, decentralized algorithms in IIoT (e.g., clustering algorithms in both wireless sensor networks (WSN) and decentralized computing) can contribute to solving the capacity and scalability issues in IIoT [28].
c) Unstable and unpredictable connections: The unstable and unpredictable connections of IIoT devices are caused not only by the mobility and sleep/idle mode of these devices, but also by unreliable wireless links to IIoT devices [29]. As a result, an IIoT network may divide into partitions that are disconnected, and these partitions may vary over time.

Based on the information of industry IoT, we present a traditional cloud-based IIoT infrastructure in the next section.

C. Cloud-based IIoT Infrastructure [30]

Fig. 2 gives an overview of a typical cloud-based IIoT infrastructure, which mainly consists of three layers: device layer, gateway layer, and cloud service layer. The device layer comprises heterogeneous IIoT devices, varying from powerful computing units to extremely low-power microcontrollers. These devices are connected to the gateway layer through various wired and wireless networking technologies, such as ZigBee, BLE, Ethernet, etc. At the gateway layer, most companies and organizations deploy their own customized gateways that manage the local IIoT networks, aggregate the data, and serve as the bridges to the clouds [31] [32]. These customized gateways are usually an integral part of the deployed IIoT infrastructure, which leads directly to “stovepipe” solutions [33]. This further causes interoperability issues; that is, data and services provided by one organization cannot be shared or utilized by devices from the other organizations (due to different networking protocols, data formats, etc.). Additionally, the employed security mechanisms are often proprietary and undocumented.

For easy understanding and presentation, we use “IoT” to represent “IIoT” in the following description.

Traditionally, the device layer and gateway layer together form the local IoT networks. A typical local IoT network consists of the following four components:

**IoT Devices:** Most IoT devices are deployed in the physical world to measure and sample their associated physical or cyber objects. They have constrained resources, including memory size, computation power, and communication bandwidth [34]. In addition, the devices and their adopted networking technologies are highly heterogeneous. This heterogeneity posts a grant challenge in interconnecting IIoT devices. It requires the interaction among the IoT devices to put the interoperability in the first place, such that heterogeneous devices are transformable in users’ acceptable forms for both syntax and semantics [35].

**IoT Storage:** In a local IoT network, a centralized storage scheme is commonly adopted for managing the IoT data, instead of either local schemes (e.g., storing data within the local memory of IoT devices) or distributed schemes (e.g., storing data within some nodes with rich storage resources in the network). In a centralized storage scheme, the data are collected by the local gateway, and then sent to and stored in local centralized storage. In our scheme, the local centralized storage could be either a historian or a private data center, in which all data are stored locally and privately. This centralized storage within a local IoT network can provide faster access to the recent data without accessing the cloud. Where and how the local storage is deployed in the local IoT network depends on the system design specification.

**Data Engine:** The data engine is a **software** component that transforms incoming and outgoing raw data to and from the IoT devices into required forms. For example, in the proposed blockchain-based IoT architecture, raw data are formed as transactions, and then encrypted and uploaded to the clouds upon request. The data engine can be deployed on the gateway or a stand-alone computing facility in the local IoT network. To guarantee the security in the local...
IoT network, the data engine also provides additional services, such as key management (e.g., distributing and updating keys to secure data transfer in local IoT network) and security mechanisms (e.g., authentication, authorization, audit services).

**Gateway:** In a typical cloud-based IoT infrastructure, the gateway is a connection entity that links the local IoT network to a cloud. On one hand, it is the sink of the local IoT network, providing data management and network management functions; on the other hand, it also serves as a P2P node on the blockchain overlay network, providing proxy functions, such as routing information provisioning, node authentication, and multicast group management [36].

In addition to the device and gateway layers, the cloud service layer provides cloud-related functionalities, such as database service and application service, for managing the data provided by the local IoT networks. Together, the local IoT networks and cloud service layer comprise the most common existing cloud-based IoT infrastructure.

### D. Issues and Challenges of Industrial IoT

In this section, we focus on the challenges of industrial IoT. For simplicity, we denote industrial IoT by IoT thereafter, without loss of generality. From Section II-C, the IoT platform guarantees the connection of various smart objects, such as sensors and actuators, which sense and collect information from the physical environment and then take some actions to react to these physical environment. The combinations of functional specifications offered by such a platform are multi-functioned, for example, device management, data analytics, cloud storage, and connectivity. While it is commonly understood that IoT technologies could play a key role as an enabler for various industrial opportunities, IoT still poses several research challenges in many aspects [5] [37].

a) **Heterogeneity:** The heterogeneity in IoT systems exhibits in several distinct aspects, for example, the heterogeneous IoT devices, the heterogeneous network topologies, and the heterogeneous IoT data types (i.e., structured, semi-structured, and non-structured). Often, the heterogeneity is the root of other challenges such as interoperability, privacy, and security.

b) **Complexity of Networks:** There are many communication/network protocols that coexist in IoT applications. It is an open issue for the standardization of IoT, which typically requires to be supported by independent and multinational governmental entities, alliances, and organizations (e.g., IEEE, IETF, W3C, IEC). These standardization processes cover many distinct aspects of IoT products, services, and systems, from communication technologies to architecture design. NFC, Bluetooth, 6LoWPAN, WirelessHART, Sigfox, LoRA, and NB-IoT are several popular network protocols in industry applications, all of which offer different network services. For example, 6LoWPAN and WirelessHART usually provide a limited range of communication (e.g., less than 100 meters), whereas LPWAN technologies can extend the coverage range from 1 km to 10 km [38] [39] [40].

c) **Poor Interoperability:** The interoperability typically refers to the capability of things in IoT systems, including both hardware and software components, to exchange, collaborate, and make use of information. The characteristics of decentralization and heterogeneity in IoT systems present a challenge for exchanging and sharing data between different industrial sectors (e.g., industrial plants of various sizes). Additionally, industrial sectors require to meet some specific capabilities for the design, implementation, without knowing the deployed solutions of other sectors, it is difficult to achieve interoperability [8].

d) **Resource Constraints of IoT devices:** Most IoT devices (e.g., sensors, actuators, RFID tags, smart meters) are resource-constrained by things such as computing resources, storage resources, bandwidth, and power supply. These devices are typically called lightweight nodes. The functionalities and applications that can be supported and deployed on these lightweight nodes are largely limited [41]. For example, it is almost impossible to deploy a complex crypto-primitive. However, without security protection, the constrained resources leave IoT devices vulnerable to various malicious attacks.

e) **Privacy Vulnerability:** Privacy is intended to ensure the appropriate use of IoT data, that is, users’ private information is not disclosed or released without the permission of the user. Preserving data privacy is challenging due to its complexity, decentralization, and heterogeneity of IoT systems. As stated in Section II-C, industrial IoT largely depends on the cloud to provide more computing and storage capabilities. However, uploading the confidential IoT data to the cloud may also compromise the vulnerable privacy of IoT [42].

f) **Security Vulnerability:** Security is an extremely important aspect for any industrial IoT application; however, the decentralization and heterogeneity of IoT systems make it more difficult to ensure security. The typical solutions for authentication, authorization, and communication encryption may not be suitable for IoT scenarios, due to the difficulty of implementing these security mechanisms on resource-constrained IoT devices. Further, IoT systems are vulnerable to various malicious attacks due to, for example, failing to update these firmware [43].

g) **Massive Data Management:** In terms of communication/transmission and storage, the volume of data generated by IoT devices can be enormous and difficult to manage. Current scalable infrastructure is not enough to handle this massive volume of data efficiently [44].

### III. BLOCKCHAIN

This section discusses some preliminary information on blockchain technologies. We focus on the technical aspects, benefits, and challenges of blockchain.

#### A. Blockchain Basics

Blockchain is a publicly known technology underlying digital cryptocurrencies, such as Bitcoin [45]. In its nutshell, the blockchain can be roughly explained as an immutable, decentralized, trusted, and shared ledger based on the underlying distributed networks (e.g., peer-to-peer (P2P)). Essentially, the blockchain is a distributed data structure, and is labeled as the “distributed ledger” in its applications functioning to record the transactions generated within network [46]. Typically, cryptocurrency is only one application of the functions of the record-keeping, and the distributed ledger technology has great potential to be adopted to other scenarios provided that the data exchanges happen. The key idea behind blockchain technology is decentralization, which means blockchain technology does not require any trusted central point or party to control or manage the participating nodes. Instead, all participating nodes (or peers) in a blockchain-enabled network maintain identical copies of its ledger. Each node has the possibility to verify other entities’ behavior within the network, as well as the capability to create, authenticate, and validate the new transactions to be recorded in a blockchain. This decentralized architecture ensures robust and secure operations on blockchain and provides various advantages (e.g., tamper-resistance and freedom of the vulnerabilities of single-point failures [47]).

---

To understand some potential applications of blockchain in industrial IoT domains, it is important to gain a basic understanding of the working principles of blockchain and how it achieves the claimed decentralization. With more transactions being executed and appended, the blockchain ledger continuously grows. When a new block is generated by a certain participating node (e.g., depending on the specified consensus protocol), it must go through the validation processes by all other nodes. Once the proposed block is validated by the majority of honest nodes, that block is automatically appended to the end of the blockchain via the inverse reference pointing to its immediately previous block. The first block of a blockchain is called the genesis block, which has no previous block. The blocks over the blockchain network achieve a distributed and decentralized synchronization via a consensus protocol, which enforces strict rules and common agreements among the participating nodes. Because the blockchain is distributed throughout the whole network, any tampering behavior can be easily detected by other nodes of the network.

1) Components of Blockchain: Forming a blockchain requires several key components, such as data block, distributed ledger, consensus algorithm, and smart contract. Fig. 3 shows a graphical representation of a generic blockchain.

a) Data Block: A blockchain consists of blocks that contain details of transactions that occurred within the network. The transaction information can be considered as a token transferring process occurring in a network or any form of data exchange. Each block can logically be divided into two major components, namely, the block header and the block body [48]. Transactions are stored in the block body, while the block header contains the metadata information of this block (e.g., the identifier of its previous block, timestamp, Merkle tree root). The blocks are then connected in a chain structure (similar to a linked list). Each block is linked to its immediately previous block via a cryptographic hash. The identifier of that block is typically obtained by taking its cryptographic hash, which is why having each block linked to its previous block helps the blockchain achieve immutability. In this way, all blocks in the chain can be traced back to only their previous one, and no chance exists for modifying or altering the appended blocks. To illustrate, for attackers to successfully alter the contents of a single block, they would have to alter the headers in all successive blocks and have this alteration taking place and getting an agreement among the majority of the nodes in the network (e.g., more than 50% of nodes), so that the peers reach a consensus on this altered blockchain. The transactions in the block body are typically arranged in a Merkle tree-based structure, where a leaf node represents a transaction submitted by a blockchain user/client. However, different applications may have different block data structures. For instance, a typical block header may contain the following essential information: (i) the previous block hash, (ii) the Merkle root storing the hash of a group of transactions in that block, and (iii) the timestamp referring to the time when the block is created.

b) Distributed Ledger: A distributed ledger is a type of database shared and replicated among the entities of a distributed network. This shared database is available and accessible for all network participants within the network. The behavior of recording transactions is similar to the process of data exchange among the participants of the network. In a decentralized setting, where no trusted third party is required to manage and control the system run, the participating nodes can automatically reach an agreement via a well-established consensus protocol. Each record associates with a unique cryptographic signature and a timestamp, which makes the ledger auditable and immutable. Any modifications on the transaction inevitably produce an altered hash within its branch, and this alteration is easily detected with little computational effort.

c) Consensus Algorithm: No centralized entities exist in a blockchain system to regulate and enforce the transaction rules or preserve data against security threats. Consensus algorithms aim to securely update the replicated shared states and ensure that all replicas of the shared states are synchronized and in agreement at any given time. A consensus algorithm in blockchain is a mechanism used to reach an agreement on a single data block between multiple unreliable nodes. For example, by solving a complex mathematical puzzle, Proof-of-Work (PoW) in Bitcoin [45] can be used as a consensus mechanism. However, the serious drawback of PoW is its high resource consumption, which would be unsustainable and unaffordable in some practical applications. Consequently, considering the practical use cases in industrial IoT, many framework designs are considering to choose the Byzantine fault-tolerance (BFT) [49] [50] protocol as a suitable candidate.

d) Smart Contracts: A smart contract is a programmable application running on blockchain, managing, and processing transactions under the specified terms and conditions. A smart contract, in practice, is a digital equivalent of a transitional economic contract between various engaging entities [51]. Unlike conventional contracts enforced by centralized authorizing entities, a blockchain network does not require authorizing intermediaries to ensure that the terms and conditions in a smart contract are met. Smart contracts have become increasingly popular in blockchain since the first smart contract platform, Ethereum [52], was released in 2015. A smart contract is sometimes termed as an “autonomous agent” or “self-executing engine”. The essence of self-execution is that once the specified conditions have been fulfilled, the codes automatically execute the contractual clauses specified by the contract. For example, in an industrial feedback control system, when an industrial process measurement is higher than a threshold defined in a smart contract, it automatically triggers an event (e.g., warning message) over the blockchain network. This triggered event is recorded as a transaction that is kept on blockchain as an immutable record. This type of self-executing agreement, relying on well-written codes, makes smart contracts unalterable and resistant to external attacks [53].

2) Types of Blockchains: Based on the way that a blockchain is used, blockchains can be classified into multiple types with some distinct attributes. In general, blockchain can be classified into three categories, namely, public (or permissionless), private (or permissioned), and consortium (or federated) blockchains [54] [55].

a) Public Blockchain: A public blockchain is an open and transparent network, which implies that anyone can join and make transactions as well as participate in the consensus process. Also referred to as permissionless blockchain, it functions in a completely distributed and decentralized way. The permissionless blockchain makes it possible for anyone to maintain a copy of the blockchain and engage in the validation process of new blocks. Typically, this type of blockchain is adopted by cryptocurrency cases, such as Bitcoin and Ethereum. A permissionless blockchain is typically designed to accommodate a large number of anonymous nodes, so minimizing potential malicious activities is essential. Due to the anonymous participating process, it requires some kind of “proofs” to show its validity of new blocks before publishing them in a public blockchain. For example, a proof could be solving the computationally intensive puzzle or staking one’s cryptocurrency. Public blockchain normally requires some kind of incentive to reward the peer nodes who attempt to publish new blocks onto the blockchain (e.g., attaching a processing fee on each submitted transaction). Public blockchain can prevent itself from being compromised by the incentive mechanism, as it would be too costly to manipulate the contents because thousands of other peers are engaged in the same decentralization consensus to validate the transactions.
Blockchain has several unique features, such as decentralization and immutability, all of which can be highly beneficial to industrial IoT applications. We briefly present these key properties \[57\] [58].

a) Decentralization: In centralized infrastructures, data exchanges (e.g., in the form of transactions) are validated and authorized by a trusted third party. Using a third party may incur a much higher maintenance cost on a centralized server and appear as a bottleneck to improve the performance. The decentralized nature of blockchain does not rely on any centralized control entity (e.g., server) for transaction handling and processing. Blockchain participating nodes follow the specified consensus protocols to validate and confirm transactions in a reliable and incorruptible manner, instead of relying on a central authority or third party to verify the transactions. This exceptional property offers some promising benefits, for example, eliminating a single point of failure, saving operational costs, and enhancing trustworthiness.

b) Immutability: The blockchain consists of a chronologically-linked chain structure of blocks in which each link is essentially an inverse hash pointing to its immediately previous block, secured by the cryptographic hash operations. Particularly, the cryptographic hashing process of a new block always contains the metadata of the hash value of the previous block, which makes the chain unalterable. All new blocks on the blockchain are agreed upon by peer nodes via a specified decentralized consensus protocol, which makes blockchain censorship-resistant and nearly impossible to be tampered with. Any modification on a block invalidates all its subsequently generated blocks. Additionally, all previously recorded data in blockchain are permanently immutable. For example, an attacker would have to compromise a majority of the participating nodes of a blockchain network to alter any previous records. Otherwise, any modification on a blockchain is easily detected.

c) Auditability and traceability: All peers in the blockchain network hold one exact copy of the chained-blocks data and can thus access and verify all timestamped transaction records. The blockchain data is essentially transparent and open to every authenticated user who can access and verify the committed transactions in a blockchain network. In other words, the same copy of records on blockchain network is easily detected.

This transparency enables peers to look up and verify transactions involving specific blockchain addresses. As described in Section III-A, it may involve different identification mechanisms (e.g., memberships) for different types of blockchains (e.g., public vs. private blockchains). For example, a public blockchain offers a privacy-preserving mechanism by pseudo-anonymity technologies, in which the chain of a block address cannot be traced back to its real owner. Typically, the industrial IoT may favor the use of either private blockchain or consortium blockchain, thus, it must provide

---

2This central authority does not participate in blockchain construction, and it mainly provides the identification-related services.
Due to their high performance and accuracy requirements in processing transactions, scaling blockchain remains a major issue in its applications (e.g., in digital finance and beyond). In industrial IoT, where a much higher volume of data transactions generates (e.g., data creation or transfer), the issues on low throughput and scalability are exacerbated.

2) Storage: The storage poses yet another critical challenge in blockchain applications. The storage issue is interconnected with the scalability issue. Although only the full nodes (the nodes that can fully validate transactions and blocks) are required to store the full chain, the storage requirements are still significant on these full nodes. As the size of the chain grows, nodes require more and more resources, thus decreasing the system’s capacity scale. Consequently, an extra-large chain has some negative effects on the system performance, such as increasing synchronization time for new users. Some research works have been proposed to deal with storage issues, such as BigChainDB [66] and Inter-Planetary File System (IPFS) [67]. For example, IPFS is a protocol designed to store decentralized and shared files to make the web safer, faster, and more open with a P2P distributed file system. IPFS aims to increase the efficiency of web services while removing duplication and tracking version history for each file.

In industrial IoT scenarios, the devices can generate a huge amount of data in a very short period, and both the data hash and the data itself need to be stored. When the chain grows over time, all participating nodes will need larger storage and higher bandwidth to keep up-to-date with the transactions added to the ledger, which may result in an increase of expensive hardware (e.g., storage disks).

3) Privacy: Protecting the privacy of users and their data records on a blockchain is a challenging task. In a basic implementation (e.g., public blockchain), data on the ledger is open to the public for verification by all miners. For example, blockchain applications in public networks (i.e., Bitcoin) have stored transactions associated with generated blockchain addresses, and all transaction records are visible to all participants of the Bitcoin network. But this also implies that any sensitive data is inherently non-private. If confidentiality is necessary for some applications, it will be required to either host a blockchain system that can be accessed only by trusted entities or to apply advanced cryptographic primitives. However, the latter option would require all miners to verify the correctness of encrypted transactions (e.g., multi-party computation and functional encryption). Still, the use of complex cryptography would limit the auditability and thus the ability to have meaningful shared governance.

Privacy in a private or consortium blockchain can be tackled differently, because by definition they must provide authentication and authorization mechanisms for all participating nodes. However, even inside a private blockchain, participants might also want to preserve the privacy of their data according to different levels of privacy.

4) Security: Due to the inherent nature of decentralization, blockchain could be vulnerable to many security threats, such as...
the 51% attacks [68]. In a 51% attack, the coordinated malicious users, by a majority (or even a large majority) of the participating nodes, can reorder, remove, and change transactions from the ledger. Blockchain applications must provide a proper incentive mechanism to keep the participating nodes working honestly. In addition, blockchain is also vulnerable to some traditional network attacks, such as Denial of Service (DoS) or partitioning attacks [69]. These attacks may be aimed at lowering the number of participating nodes or fracturing network nodes to prevent the consensus protocol, thus lowering the bar for 51% attacks or creating some inconsistent states.

Additionally, smart contracts often exploit some loopholes. For example, an adversary exploiting the shortcoming of a smart contract was seen in the DAO attack\(^3\) [70] [71]. Thus, when developing security standards, it is critical to script smart contracts in such a way that no loopholes exist that may compromise the security of the devices in IIoT networks.

### IV. Integration of Blockchain and IIoT

Blockchain establishes on a decentralized network (e.g., P2P network) that reduces the cost of installation and maintenance in centralized infrastructures (e.g., data centers), and reduces the cost of networking equipment by distributing computational and storage requirements among all devices. In general, the decentralized communication model eliminates the issue of the single point of failure in a traditional centralized network. Moreover, the decentralized model has been adopted in many decentralized industrial infrastructures. By integrating tamper-resistant ledgers [57], the decentralized model can achieve many features desired to the IIoT networks, such as reliability and interoperability. However, it is a challenging task to apply the existing blockchain technologies directly in resource-constrained industrial IoT networks.

This section discusses the integration of blockchain and industrial IoT, including the motivation, the basics on blockchain-enabled IIoT platforms, and some fundamental blockchain-enabled technologies.

#### A. Motivation

The blockchain shows great potential in overcoming the interoperability issues in IIoT. According to the International Data Corporation, up to 20% of IoT deployments will offer blockchain-enabled services, and more than 10% of global GDP will be related to the blockchain-enabled systems by 2027 [72].

1) **Blockchain will Revolutionize IIoT:** In industry, the inexpensive concept of blockchain-based data-keeping and accounting will initiate various innovative technologies that will encourage enterprises to create their tamper-resistant ledgers and accounting systems. This will potentially revolutionize the respective industries in general. Most existing IIoT solutions are based on a centralized server-client model that is connected via the Internet to cloud servers. While these solutions are sufficient for today’s applications, with the advancement and extension of IIoT networks, tremendous demands will come for new solutions and platforms that make networks more decentralized [73]. The development of large-scale decentralized networks (e.g., P2P networks) is one of those potential solutions. Generally, blockchain can offer a safer and more reliable option for enterprises and individuals to ensure trustworthiness and immutability among the participants. For example, blockchain allows collaborative companies to reliably maintain and record shipping records across multiple entities in a supply chain. The blockchain can improve the interoperability in logistics by allowing supply change to operate more efficiently and more reliably.

In industrial applications, such as an industrial control system (ICS), trustworthiness is a major challenge [47]. The nature of trustworthiness in blockchain can offer ICS a much safer environment, in which blockchain establishes a broad range of cybersecurity opportunities that would affect entire industrial systems. For instance, blockchain can ensure the entire system secure and irreversible. Typically, IIoT is an extensive network that integrates a huge number of devices, so IIoT faces various vulnerabilities and attacks. As the number of new devices connected to IIoT increases, the vulnerabilities will increase exponentially because each device cannot be guaranteed to function securely and honestly. To speed up processing performance, such as for applications requiring a real-time response, many IIoT platforms adopt lightweight solutions; for example, solutions that do not involve robust crypto-primitives to guarantee security. However, some cryptographic algorithms or primitives (e.g., SHA1) have a limited lifetime before they break, which means the current secure algorithms can get compromised if the hackers adopt and learn more advanced hacking technologies [74]. Thus, we need technologies (e.g., fault-tolerant technology) to ensure that even if parts of a system were compromised, the overall system would remain safe and secure. These technologies can be well complemented by blockchain technologies.

Blockchain will revolutionize the IIoT technologies. On one hand, in IIoT, the decentralized nature of blockchain technology will play a key role between two untrusted devices to keep devices information about their interactions, state, and digest of exchanged data. On the other hand, blockchain can significantly reduce the risks that the users are currently facing, and save the cost of business processes.

2) **How Blockchain Supports Industry 4.0:** As a new technology, the blockchain introduces new features to both industrial IoT and Industry 4.0. The communication, interaction, and commodity exchanges among the industrial sectors are built on trustworthiness. With more collaboration occurring among industries, a transparent, democratic, decentralized, efficient, and secure architecture is needed to create a trustworthy environment. Although Internet communication was possible decades ago, it still could not provide a built-in trust that is highly expected for business communications. The advent of blockchain allows people to conduct trade with anyone, even without a prior relationship or trusted third party, which undoubtedly refines the entire structure of business models in industries. Depending on different use cases, blockchain can establish various hybrid models, for example, private blockchain, public blockchain, or federated/consortium blockchain. Typically, the industrial use cases employ either private or consortium blockchains, which require permission to access these data and networks.

The impact of blockchain on industrial IoT and Industry 4.0 will be enormous. With blockchain technology, uncertainty will be eliminated, and transparency will be instilled among industrial sectors. For instance, many industrial sectors face problems in managing multiple vendors in a horizontal supply chain system, where each vendor has its own individual policy and architecture. This issue creates a communication barrier between different industrial sectors. With blockchain, even without a central authority, each industrial sector can independently track, monitor, and validate other sectors’ activities. From the raw materials to the completion of the product life-cycle, the whole cycle is not only open to the stakeholders of the industry but also to other collaborators. For each participant involved in the product development life-cycle, blockchain offers vertical networking for smart production systems. Fig. 4 shows that blockchain and industrial IoT suit each other in several aspects. This suitability makes blockchain a serious contender for becoming a member of industrial IoT and Industry 4.0.

Blockchain provides a new platform for the digital information

---

\(^3\)DAO is short for “Distributed Autonomous Organization”. 

transformation of current industries to adapt to Industry 4.0. It offers some unique features to current industrial architectures, and makes them more acceptable and dynamic. The left side of Fig. 4 shows the needs of industry, in which autonomy, CPS, P2P communication, and other latest technologies are the necessities to build an illusion of modern industry. The right side of this figure represents the characteristics that blockchain can provide, such as scalability, transparency, decentralization, and secure communication. These unique characteristics are capable to fulfill the requirements of Industry 4.0.  

3) Blockchain Platforms for IIoT: Typically, in an industrial use case, the blockchain platform plays a critical role in delivering connected operations and assets, as well as enabling some unique properties, including connectivity, big data analysis, and application development. Most existing industrial facilities, such as micro-grids, smart-grid IoT, or vehicular ad hoc networks (VANETs), are unable to connect to the IIoT with their built-in intelligence; thus, various interfaces are required to communicate with IIoT. Most, if not all, current blockchain technologies in IIoT applications focus on the design of the application layers; these underlying networks are typically abstracted to be a P2P connection without physical restrictions on network topologies, devices, and communication bandwidth.

In the IIoT domain, smart IoT devices can utilize the existing crypto-currencies-related techniques (e.g., gas in Ethereum) as an incentive scheme to record and exchange transactional activities within the network. For example, the Ethereum Virtual Machine (EVM) [75] platforms are extensively used in IoT; these platforms have built-in smart contract functionality and flexible consensus strategy, where the specified smart contract provides down-compatibility to the IIoT applications. The Hyperledger series (e.g., Hyperledger Fabric [76]) is another popular open-source blockchain platform developed by IBM; this platform offers distributed industrial components with consensus and membership strategies, and can well support by the IBM Watson IoT platform. The Hyperledger has great potentials to speed up the IIoT applications [77]. Additionally, blockchain can provide a service layer [76] [78] [79] when integrating with typical IoT architectures. For instance, Enigma, which is a blockchain-based on P2P network for decentralized personal data management, can serve as a service layer to the underlying applications [80] [81].

More blockchain platforms are being developed (e.g., Multichain [77], Litecoin [82], Quorum [83], and SMChain [84]), which provide IIoT applications with some new features, such as traceability and trustworthiness. The performance of these blockchain platforms can be measured using various metrics, such as energy consumption, CPU utilization, memory utilization, the size of the block, and so on.

In addition to the traditional chain structure, a specific platform called IOTA [85] aims at providing blockchain-like solutions specifically for IoT networks. IOTA is developed based on the technology, “Tangle,” which is designed with no chains, no blocks, and no fees. Instead of using the chain structure, Tangle inherits the anti-tampering, decentralized blockchain ledger using a directed acyclic graph (DAG) structure. The key idea is that transactions are IOTA’s only storage units, and each transaction need to confirm two more transactions that were previously published. This mechanism is much similar to the PoW scheme, which requires the participating nodes to contribute to the Tangle’s construction.

With more mature technologies on both blockchain and IIoT applications, more and more sophisticated and professional blockchain-based IIoT platforms will emerge to fit these specific application domains.

B. The Architecture of Integration

This section discusses the potential architectures that can be used to integrate blockchain into IIoT platforms. Then we discuss the main features of blockchain-enabled IoT.

In general, the blockchain nodes in an industrial scenario can be roughly classified into two types: full nodes and lightweight nodes [86]. The full nodes typically require downloading and checking all blocks and transactions in the chain, and these nodes can serve as mining nodes and can create blocks for blockchain. In contrast, the lightweight nodes typically have limited resources; they can store and process only a small amount of data for blockchain. The lightweight nodes (e.g., smart devices or sensors) generate new transactions (not blocks) that are propagated between the full nodes and eventually add the newly generated blocks via a consensus process into the blockchain.

1) System Architecture: Integrating blockchain into IIoT platforms can enable the automatic communication of the devices, which might be untrusted ones, in a distributed and verifiable manner. Section II-C presents a traditional cloud-based IIoT platform; while integrating blockchain into the IIoT platform, the blockchain layer can be considered as a middleware between the communication layer and the industrial applications. Blockchain as a middleware offers some advantages: 1) providing abstraction from lower layers of IIoT platforms, and 2) providing users with the blockchain-based services [87]. The blockchain middleware typically organizes as a composite layer, which has the potential to hide the heterogeneity of lower layers (e.g., the communication technologies of IIoT platforms [88] [89]). In particular, to support various industrial applications, the blockchain middleware layer offers various blockchain-based services that typically are implemented as Application Programming Interfaces (APIs). Fig. 5 shows an overview of the integration of blockchain into an IIoT platform. Specially, the blockchain middleware layer consists of five sub-layers [90]. Note that different IIoT applications may have different system architectures for these integrations. Here we provide a generic platform for the integration of blockchain into industrial platforms.

a) Blockchain Data Layer: This layer focuses on data collection and processing schemes. The blockchain data layer collects IIoT data from the lower layers (e.g., the perception layer) and performs certain basic data processing operations, such as encrypting data with digital signature via asymmetric cryptographic algorithms and hash functions. For instance, after the distributed validation on the consensus nodes, these consecutively connected data blocks are used to construct...
the blockchain. Different blockchain platforms may choose distinct cryptographic algorithms and hash functions to construct these blocks, which are then sent to the network layer for propagation. Data block, chain structure, Merkle tree, hash function, cryptographic algorithms, and digital signature are the major components for this layer [90].

b) Blockchain Network Layer: This layer mainly provides the required network services. The blockchain network layer is essentially a decentralized overlay network (e.g., P2P network) running on top of the communication layer [91]. The overlay network links the participating nodes together, either virtually or physically, based on its underlying communication network (e.g., wired / wireless). This layer typically records all data in a decentralized and private manner. For instance, a node may simply broadcast the transaction blocks to its connected peers; once receiving these blocks, other peers are required to verify the validity of received blocks locally. Only the valid blocks can be further propagated to other nodes over the overlay network. Propagation protocol, overlay routing, and verification mechanism are the major components in this layer.

c) Blockchain Consensus Layer: This layer mainly provides a consensus service to get an agreement on blocks among the distributed and decentralized participating nodes. The blockchain consensus layer basically involves a specified distributed and decentralized consensus protocol to build the trustworthiness of blockchain [90]. The consensus layer can use various consensus protocols, such as PoW, PoS, BFT-related protocols (see Section IV-D for details), to establish an agreement. It is worth mentioning that the block propagation mechanism (e.g., P2P relay network propagation or gossip protocol [92]) is the prerequisite for a distributed consensus protocol. Typically, for industrial use cases, the consensus protocols focus on the BFT-related consensus protocols to achieve an instantaneous agreement on data records and instant finality on data blocks.

d) Blockchain Incentive Layer: This layer typically is an optional layer, which provides a reward or incentive mechanism to the participating nodes for these efforts on consensus processes. The blockchain incentive layer is responsible for incentive-related tasks, e.g., designing a fair reward mechanism, issuing and distributing digital currency or tokens, and handling transaction costs. In particular, it is critical to design an appropriate and fair incentive mechanism for distributing the rewards to the participants who contribute to the distributed consensus. This is extremely important, especially for consortium blockchains, in which multiple organizations collectively build a blockchain and they all need to get a fair-share of the reward. Currency issues and distribution mechanisms, reward mechanisms, and transaction costs are the major components in this layer.

e) Blockchain Service Layer: This layer provides system interfaces between the components of the IIoT platform and the blockchain. The blockchain service layer provides clients with blockchain-based services for various industrial sectors, including manufacturing, logistics, supply chains, food industries, and utilities. The blockchain can be applied through smart contracts as a service (or the term “BaaS”). These smart contracts specify the rules around an agreement that lets blockchain participants exchange information, resources, and shares in a conflict-free way, while avoiding a middleman’s service (e.g., a third verification party). These smart contracts can be activated when a special event, defined by smart contracts, occurs.

It is worth mentioning that the blockchain network layer that is developed on top of the communication layer is an abstraction of underneath communication network. It can offer universal network access across multiple distinct industrial networks.

The realistic deployments of the integration of blockchain and IIoT platforms are of great importance. Due to the feature of resource constraints on lightweight nodes, storing the whole blockchain at these nodes is impossible. An IIoT network consists primarily of lightweight nodes (e.g., smart sensors, RFID readers, smart meters) and a small amount of powerful full nodes (e.g., data analysis servers, edge computing servers). In practice, a full node (alternatively called consensus node) can be a cloud server or an edge server with adequate computing resources, having a large storage space to save the entire blockchain. In IIoT, the lightweight nodes can connect peers running as full nodes to send and receive transactions. And these lightweight nodes can store only minimal blockchain information (e.g., the latest block information) but can send output requests, via messages, encoded in the deployed application protocols.
Then, the full node sends back a response that can be verified by a lightweight node by only checking its token (e.g., data and states). If passed, the lightweight node proceeds to construct the transactions; otherwise, the lightweight node returns an invalid response with modified output. Fig. 6 shows a possible deployment scenario of this integration, in which the full nodes (cloud server and edge server) store the whole/partial blockchain, while industrial devices store only the partial blockchain data.

It is worth mentioning that the lightweight nodes highly rely on the connection to the full nodes. In an IIoT environment, a lightweight node can establish connections with multiple untrusted full nodes to support output retrieval, proof generations, updates to the structure, and conflict resolution. In reality, there may be several possible ways to integrate IIoT and blockchain. More sophisticated mechanisms need to be built in order to improve the protection between different protocols in different industrial sectors.

Fig. 6. Lightweight node data retrieval from multiple full nodes.

2) Opportunities of Integrating Blockchain with IoT: As discussed in Section II-D, industrial IoT systems are facing many challenges, such as heterogeneity, poor interoperability, and resource constraints on devices. Blockchain technologies can complement the current IIoT platforms to resolve these challenges. Integrating blockchain into IIoT platforms provides several potential advantages over traditional IIoT platforms.

   a) Enhanced Interoperability: Interoperability is a big challenge in many IIoT applications. In most existing IIoT platforms, interoperability is managed at the application level, where the operators are demanded to be proficient in various (or even completely different) operations. Moreover, a huge amount of data will be generated from interconnected facilities of different IIoT applications, requiring a high degree of interoperability. Most existing Operational Technology (OT) systems typically operate in separate states, which unavoidably increases the running cost and complexity of the practical IIoT deployments. It is always a challenging task to bridge the gaps of the shared data between smart facilities from various manufacturers (or even within an organization).

   By transforming and storing data records into a shared blockchain, blockchain can potentially improve the interoperability of IIoT platforms. This process will seamlessly establish the connections between assets and information operating in different data protocols. For example, industrial sensor measured data can be translated into commonly used JSON or XML formats [93]. During this procedure, heterogeneous IIoT data types are converted, processed, extracted, compressed, and finally stored into a blockchain. Besides, the interoperability exhibits in readily passing through various types of fragmented sub-networks since blockchains are established on top of

Fig. 7. Interoperability of blockchain-enabled IIoT architecture.

the decentralized overlay network which supports universal Internet access.

On the other hand, interoperability in IIoT scenarios also refers to the process of data exchanges between different entities, for example, multiple companies following the same standard. In this scenario, the interoperability then refers to the ability of different IIoT platforms and applications to communicate, collaborate, exchange data, and use the exchanged information [94]. It can potentially reduce the duplicated information and improve system efficiency, which is essential to reducing the production cost. As shown in Fig. 7, the blockchain-enabled IIoT platforms authenticate the authorized users directly retrieve the data from the platform, in which a user can authorize the data sharing process between two distinct platforms (e.g., using a built-in smart contract) without resorting to a formal business relationship. To obtain interoperability, the blockchain-enabled IIoT platforms must store some public information of the corresponding counterparts, such as authorization rules, user-associated public keys, and data access audit logs. This can significantly improve the interoperability in IIoT applications.

   b) Improved Security: Blockchain can provide safety-enhancing solutions through important security features, such as confidentiality and availability that are inherent in the blockchain. It will secure the IIoT data, as all valid records are stored as blockchain transactions that are encrypted and digitally signed by some crypto-primitives (e.g., elliptic curve digital signatures [95]). This process ensures that all interactions with the IIoT platform remain confidential under blockchain-enabled signatures. In addition, with the decentralization feature inherent in blockchain, data is replicated across all network members without single failure bottlenecks, thus promising to provide enhanced availability. Combined with traditional cloud-based IIoT architectures, the resourceful cloud can provide off-chain storage solutions to support data availability of the on-chain storage mechanisms; even then, the IIoT network is interrupted due to external attacks. Moreover, implementing the blockchain on the cloud-based IIoT platforms may enhance the security of the blockchain system itself. For example, clouds can use their available and powerful network security tools to maintain and preserve blockchain software (i.e., mining mechanism), against potential threats.
The blockchain offers a secure, immutable, and trustworthy platform, which can tolerate a sufficiently large network even with untrusted peers. The privacy of the blockchain-enabled IIoT platforms can also be enhanced by data ownership, data transparency and audibility, and fine-grained access controls.

c) Greater Transparency: Blockchain technology enhances the transparency of data exchange and transactional data exchange. As a distributed and decentralized ledger, all participants in the network share the same information in their individual copies, which can only be updated via consensus protocols. Additionally, these updates must be agreed upon by every participant. Any modifications on a single transaction would require the alteration of all subsequent records and potentially require collusion over the entire network [96]. The data on the blockchain is available and accessible for all authenticated participants, and is more accurate, consistent, and transparent than traditional IIoT platforms (e.g., without blockchain).

d) Improved Traceability: The blockchain provides the possibilities to solve important glitches in traceability challenges that occur on traditional IIoT platforms. Due to the heterogeneity of IIoT devices and the complexity of interactions between the information providers, it is quite challenging to accurately track and link information to the content without any biases between different providers. Traceability is very important in the verification of industrial transactions among different industrial sectors. Typically, blockchain data is required to be identified and verified anywhere and anytime. All transactions stored in blockchain should be traceable. For example, Lu et al. [97] creates a blockchain-based platform for product traceability, which offers traceable services to both suppliers and retailers. Tractability makes it possible to inspect and verify the quality and originality of the goods at any stage of a product life cycle. Moreover, the feature of immutability assures the reliability of data because it is nearly impossible to modify or falsify any transactions that are already stored in the blockchain.

e) Improved Corporation: Integrating blockchain into a traditional cloud-based IIoT platforms, particularly with multi-clouds scenarios, is a promising research topic. Integrating blockchain enables boundlessly corporate cloud service providers with IIoT users, without the requirement of a central authority. The IIoT data is securely transmitted under blockchain management, even in an untrusted environment. User anonymity can also be ensured, as blockchain can hide users’ sensitive information (e.g., via pseudo-anonymous identity) to avoid potential data leakage issues. Especially, the use of smart contracts in blockchain allows for a secure data sharing process in cooperative cloud-based IIoT networks by offering automatic user authentication and data access capabilities without trusting any third parties. It potentially improves the corporation in cloud-based IIoT, paving the way to feature large-scale IIoT applications.

f) Autonomic Interaction: Via smart contracts, blockchain can grant automatic communication between IIoT devices or subsystems. For example, Distributed Autonomous Corporations (DACs) aim to provide automated transaction services in which traditional roles, like governments or companies, are not involved with the transactions [98]. As they are implemented by smart contracts, DACs can work automatically without human intervention, consequently saving on operational costs.

g) Reduced System Complexity: Integrating blockchain with cloud-based IIoT platforms can significantly reduce the complexity of system implementations. This integration is known as Blockchain-as-a-Service (BaaS) [99], where all established blockchain platforms are available to set up and run blockchain for industrial applications without worrying about the underlying hardware technologies and infrastructures [100]. In addition, blockchain platforms can be run online using cloud infrastructures, which aims to reduce resource overheads for running blockchain in resource constraints IIoT devices. The integration of blockchain and cloud-based IIoT opens up numerous opportunities, with simple and cheap applications, for accelerating large-scale industrial IoT deployments.

C. Identification and Data Structure

In recent years, both the blockchain and IIoT technologies have gained great attention in many industrial applications, including supply chains, logistics, manufacturing, and smart grids. This section discusses some key components in blockchain-enabled IIoT platforms.

1) Identification and Tracking Technologies: In IIoT, devices are not isolated and typically have relationships with other devices. Additionally, their ownerships are subject to change. Identity management involves the processes related to both authentication and authorization, and prevents any malicious use without the access privileges. Classic authentication schemes, such as user ID and password combinations, often do not work well in IIoT scenarios because users are not actively involved and devices automatically authenticate themselves using tokens or security certificate mechanisms. Certain security mechanisms should have been put in place on the implementation of IIoT to prevent the abuse of identities. There are also many identification management platforms (e.g., OAuth [101]) providing an open authorization framework. The common issue with the conventional identity management approaches, however, is the lack of assured trust and reliance on approving authorities from the trusted third parties. Also, interoperability is an ongoing challenge in the presence of multiple protocol options, cross-platform architectures, and variations in semantics and conformance [8].

A blockchain-enabled IIoT platform requires the identification information to be provided to every device; this information is used to identify all transactions a device published. There is a lot of research on managing the identities of large-scale connected devices in a decentralized IoT platform. For example, Axon et al. [102] highlights the potential benefits of the PKI without single points of failure by using blockchain, which demonstrates variable levels of privacy-awareness that can be achieved without blockchain-based PKI. In a private blockchain, the peers also need to be authorized before entering a blockchain network. For example, Hyperledger Fabric provides identity management to implement the enrollment and transaction certificates [76].

In general, identification technologies are critical in providing authentication, authorization, and access control services in any kind of IIoT platform. Several key technologies also exist that together manage and track identifications [103]: 1) device identification in IIoT platform, which includes the pseudo-identity generation for IIoT devices, users, and services using public-key-based pseudo identities generation; 2) communication technologies, in which the machine-to-machine communication is the mainstream; 3) networks technologies, which include 5G, mobile networks, and industrial sensor networks.

Blockchain provides a shared and immutable ledger, which every authenticated user can access and use to track the recorded transactions [104]. This ledger potentially enhances the data’s trackability among IIoT platforms. Every individual node has an exact copy of transactions in the form of the block, thus, the track becomes much easier. While considering privacy, the trackability focuses mostly on the verification process of transactions, which is used to verify if a transaction is indeed generated.

Blockchain-based identity and access management systems can be leveraged to enhance IIoT security. These systems have already been used to securely store information on provenance, identity, credentials, and digital rights of things. Provided that the original
### TABLE II. DESCRIPTION OF A TRANSACTION

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>From</td>
<td>The address of local metering device, e.g., UUID of meters</td>
</tr>
<tr>
<td>To</td>
<td>The target gateway, either field gateway or edge gateway, that the metering measurement is sent to</td>
</tr>
<tr>
<td>Type</td>
<td>What type of measurement, e.g., warning</td>
</tr>
<tr>
<td>Device_info</td>
<td>The information of metering device</td>
</tr>
<tr>
<td>One_Time_PK</td>
<td>The device’s one-time public key used to encrypt the message from device to gateway so gateway can verify its integrity and confidentiality</td>
</tr>
<tr>
<td>TimeStamp</td>
<td>Unix timestamp when a device is measured its measurement (assuming all plants are) synchronous locally. Also, a timestamp is used to accept as valid if it is greater than the timestamp from the previous data block</td>
</tr>
<tr>
<td>TX_ID</td>
<td>To identify the order of measurement from “from” to the same “to”. Each measurement has a unique ID during its block epoch</td>
</tr>
<tr>
<td>Data</td>
<td>Measured value from physical devices</td>
</tr>
<tr>
<td>Hash_Type</td>
<td>Indicate what digest algorithm used, e.g., SHA-256, SHA-512</td>
</tr>
<tr>
<td>TX_Hash</td>
<td>The digest of the measured value</td>
</tr>
<tr>
<td>Sig_Type</td>
<td>Indicate what signature algorithm used</td>
</tr>
<tr>
<td>Signature</td>
<td>The signature of the measurement</td>
</tr>
</tbody>
</table>

### TABLE III. DESCRIPTION OF A DATA BLOCK

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Block Header</td>
<td></td>
</tr>
<tr>
<td>Hash_Pre_Data_Blk</td>
<td>Hash of previous data block. Each data blk is inherited from its previous data block, since it uses the previous block’s hash to create the new block’s hash.</td>
</tr>
<tr>
<td>Block Hash</td>
<td>An identifier to identify a block, which is a cryptographic hash.</td>
</tr>
<tr>
<td>Version</td>
<td>The block version number, with which the system can upgrade the software and specify a new version.</td>
</tr>
<tr>
<td>Merkle Root of TXs</td>
<td>Merkle tree root, a data structure that summarizes the transactions in the block.</td>
</tr>
<tr>
<td>No. of TXs</td>
<td>Identify the number of transactions to be included in block body.</td>
</tr>
<tr>
<td>Signature</td>
<td>The signature of the block, which is signed by the creator of the block.</td>
</tr>
<tr>
<td>Timestamp</td>
<td>Show the time when a new block created.</td>
</tr>
<tr>
<td>Data Block Body</td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Shows the order of transactions in one data block sequentially from 1 to N, where N is the total number of TXs in this block.</td>
</tr>
<tr>
<td>TX_ID</td>
<td>Extracted from Transaction.</td>
</tr>
<tr>
<td>TX Data</td>
<td>Extracted from Transaction.</td>
</tr>
<tr>
<td>TX Hash</td>
<td>Extracted from Transaction.</td>
</tr>
</tbody>
</table>

### a) Industrial Transactions

Transactions in cryptocurrencies (e.g., UTXOs in Bitcoin [45]) are quite different from industrial transactions, as they need to carry the industrial information on their own transactions. In the following description, we use a smart metering system as an example to outline the basic structure of an industrial transaction, which can be generalized into other industrial cases. Table II shows a conceptual structure of the transaction with description.

### b) Industrial Block

A block in our industrial blockchain is called a “data block”. A data block is directly related to the transactions, which come from physical resources and local networks. Each data block consists of two parts: a block header and a block body. The header contains metadata about its block. The body of the data block contains the transactions. These transactions are hashed only indirectly through the Merkle root. The description of each field of a data block is as shown in Table III. Notice that most cryptocurrencies (e.g., Bitcoin) store only the transactions’ hashes and the Merkle tree root into the blockchain, while industrial cases need the whole transaction to be stored in the data block for further analysis in condition monitoring.

### D. Consensus Classification in Blockchain

This section presents the state-of-the-art consensus protocols for blockchain protocols in a general way [62]. These protocols can be further adopted to blockchain-enabled IIoT platforms.

In general, the consensus protocols can be put in two categories when being used in the blockchain: PoX and BFT. We know Proof-of-Work (PoW) mechanism on Bitcoin [45] and Proof-of-Stake (PoS) on Ethereum [105]. Technically speaking, PoW and PoS are not the decent “consensus protocol”, whose mechanisms are used for
determining the membership or the stake in a Sybil-attack-resistant fashion. Due to historical reasons, (e.g., Bitcoin used PoW as a “consensus” protocol to build a bitcoin blockchain), we literally categorize them into consensus protocols. For example, in a hybrid consensus (e.g., ByzCoin [106] and Hybrid Consensus [107]), the decent consensus protocol (the algorithm for agreement on a shared history) is separable from and orthogonal to the membership Sybil-resistance scheme (e.g., PoW). Here, we use Proof-of-X (PoX) is used to represent all alternatives of proof-of-something (including PoW and PoS), and use BFT is used to represent Byzantine-based consensus protocols. In industrial applications, both PoX and BFT work together to achieve the consensus process.

1) PoX: Most PoX-based consensus protocols require that the participating node has some efforts or resources in place to prove its validity as a miner. We take PoW and PoS as examples to illustrate the PoX mechanisms.

In blockchain, PoW is also called Nakamoto consensus, named after its originator [108]. Nakamoto was proposed in 1992 for spam Email protection. In PoW, the nodes that generate hashes are called miners and the process is referred to as mining. When applying PoW as a general consensus in blockchain, it is subject to various kinds of attacks [45], such as forks, double-spending attacks, and 51% attacks. These are the general problems in PoW consensus. However, when implementing PoW into blockchain protocols, due to running PoW locally, special care is required (e.g., selfish mining [109]). Selfish mining allows colluding miners to generate more valid blocks than their computing power would normally allow if they were following the standard protocol. These valid blocks are typically generated ahead of time, so that the colluding miners withhold blocks that they have found, and then select a favorite one to maximize these advantages (e.g., controlling one shard). Thus, applying PoW into blockchain requires agreed epoch randomness for each epoch.

Compared to PoW, PoS protocols replace wasteful computations with useful “work” derived from the alternative commonly accessible resources. For example, participants of PoS vote on new blocks weighted by their in-band investment, such as the amount of currency held in the PPcoin blockchain [110]. In general, PoS has a candidate pool which contains all qualified participants, called stakeholders (e.g., the amount of stake is larger than a threshold value) [111] [112]. A common approach is to randomly elect a leader from the stakeholders, which then appends a block to the blockchain. However, in blockchain, PoS could be subject to grinding attacks [113], in which a miner re-creates a block multiple times until it is likely that the miner can create a second block shortly afterward. It should be mentioned that PoS is not just one protocol, but instead a collection of protocols. Many PoS alternatives exist, such as Algorand [114], Ouroboros [105], Ouroboros Praos [112], Ethereum [115].

In addition to the main PoS protocol, other PoX-based alternatives exist which require miners to hold or prove the ownership of assets. We list three alternatives: proof-of-deposit (PoD) [116], proof-of-burn (PoB) [117] and proof-of-coin-age (PoCA) [118]. Readers are referred to the corresponding papers for details about these alternatives.

2) BFT: Most practical blockchain systems use classic BFT consensus protocols, for example, PBFT in industrial cases. In this section, we focus on the potential BFT consensus protocols in blockchains, or their novel compositions that can be tailored for use as the consensus protocols in blockchains. Roughly speaking, BFT protocols can be classified into two categories; leader-based BFT and leaderless BFT. Most BFT protocols are leader-based, for example, PBFT or BFT-SMaRt [119]; leaderless protocols include SINTRA [120] and HoneyBadger [121].

Actual systems that implement PBFT or its variants are much harder to find than systems that implement Paxos/VSR [122]. BFT-SMaRt [123], launched around 2015, is a widely tested implementation of BFT consensus protocols. Similar to Paxos/VSR, Byzantine consensus, such as PBFT and BFT-SMaRt, expects an eventually synchronous network to make progress. Without this assumption, only randomized protocols for Byzantine consensus are possible (e.g., SINTRA, which relies on distributed cryptography [120] and HoneyBadger [121], which can achieve eventual consensus on an asynchronous network).

Still, many well-known blockchain projects use PBFT and BFT-SMaRt protocols. For example, Hyperledger Fabric [76] and Tendermint Core [124] implement PBFT as their consensus protocols; Sym- biont [125] and R3 Corda [126] use BFT-SMaRt as their consensus protocols. We briefly discuss these two leader-based BFT consensus protocols, which can be used as intra-shard consensus process.

a) PBFT: PBFT can tolerate up to 1/3 Byzantine faults. We brieﬂy describe its consensus procedures. One replica, the primary/leader replica, decides the order for clients’ requests and forwards them to other replicas, the secondary replicas. All replicas together then run a three-phase (pre-prepare/prepare/commit) agreement protocol to agree on the order of requests. Each replica processes every request and sends a response to the corresponding client. The PBFT protocol has the important guarantee that safety is maintained even during periods of timing violations; progress only depends on the leader. On detecting that the leader replica is faulty through the consensus procedure, the other replicas trigger a view-change protocol to select a new leader. The leader-based protocol works very well in practice and is suitable in blockchain, however, it is subject to scalability issues.

b) BFT-SMaRt: BFT-SMaRt implements a BFT total-order multicast protocol for the replication layer of coordination service [119]. It assumes a similar system model to BFT SMR [127] [128]: \( n > 3f + 1 \) replicas to tolerate \( f \) Byzantine faults, unbounded number of faulty-prone clients, and eventual synchrony to ensure liveness. Typically, the BFT-SMaRt consists of three key components: Total Order Multicast [129], State Transfer [130], and Reconfiguration [131]. Refer to [129], [130], [131] for further details.

In addition to the above legacy leader-based BFT protocols and the mentioned BFT protocols, several variants or newly invented algorithms exist (e.g., Hotstuff [132], Tendermint [124], and Ouroboros-BFT [133]). Refer to the corresponding references for details.

We now brieﬂy discuss the leaderless BFT protocols. These types of BFT protocols mainly target the asynchronous settings, which are based on randomized atomic broadcast protocols. Unlike existing weakly/partially synchronous protocols, in an asynchronous network, messages are eventually delivered but no other timing assumption is made. We take SINTRA [120] and HoneyBadger [121] as examples to describe the leaderless BFT protocols.

c) SINTRA [120]: SINTRA refers to a Secure INtrusion-Tolerant Replication Architecture for coordination among the large-scale participating nodes in the asynchronous setting, which also is subject to Byzantine faults. SINTRA presents a new asynchronous atomic broadcast protocol [134], which includes a reduction mechanism to simplify the atomic broadcast (ABC) protocol to a common subset agreement (ACS), and this greatly improves the performance. By utilizing the threshold crypto- primitives (e.g., threshold signature), the security is further enhanced.

d) HoneyBadger [121]: HoneyBadgerBFT essentially follows asynchronous secure computing with optimal resilience [135], which uses reliable broadcast (RBC) and asynchronous binary Byzantine agreement (ABA) to achieve ACS. HoneyBadger cherry-picks a bandwidth-efficient, erasure-code RBC (AVID broadcast) [136] and the most efficient ABC to realize. Specifically, HoneyBadger uses a threshold signature to provide common coins for randomized ABA
protocol, which achieves higher throughput by aggressively batching client transactions.

Besides the above two leaderless BFT protocols, some other peer-reviewed and non-peer-reviewed works exist, such as HotStuff [132] and DBFT [137].

E. Smart Contracts on Blockchain

Nick Szabo introduced the concept of a smart contract in 1994, defining a smart contract as “a computerized transaction protocol that executes the terms of a contract” [138]. The interaction is mediated by smart contracts in blockchain-enabled IIoT platforms, where smart contracts can encode and drive the business logic processes. For an IIoT platform, the smart contract can be implemented in a more efficient and reliable decentralized manner. Within the blockchain domain, smart contracts are the scripts stored and executed on the blockchain. As smart contracts can reside on the chain, they have a unique address to identify which contracts they target (e.g., different versions on one contract). In blockchain domains, smart contracts perform the functionality by carrying out transactions in a predetermined fashion, agreed upon by parties participating in the contract.

The smart contract can help the participants in a blockchain system exchange data, assets, and shares in a conflict-free way, thus avoiding the middleman services [139]. Essentially, there are several key components in a smart contract: parties, triggering events, and regulations. A smart contract can be triggered by addressing a transaction to it. It then executes automatically and independently in a prescribed manner on each node of the blockchain network, according to the data that was included in a triggering transaction. Triggering events in a smart contract generally incur an execution fee, as an invocation itself is considered to be a valid transaction that will show the intention to be recorded into a blockchain. Execution fees incentivize peers to publish new blocks and mitigate the network’s flooding attacks.

A blockchain that supports smart contracts enables a multi-step process or interaction between the counterparties that might be mutual. In general, the transacting entities must perform several tasks: 1) inspect the code and identify its outcomes before making the decisions to participate with the contract; 2) ensure the execution, since the code is already deployed on the network that neither of them controls fully; and 3) verify the process, since all interactions are digitally signed. The possibility of disagreement is eliminated when all possible outcomes are accounted for, because the participants cannot disagree over the final outcome of this verifiable process in which they are engaged. Smart contracts typically operate as autonomous and independent agents whose behaviors are completely predictable, as they can be trusted to push forward any on-chain logic.

The smart contract can be used to perform a variety of functions within a blockchain network. The following list shows several practical functions in IIoT networks.

1. Allowing the “multi-signature” transactions, where a transaction is only carried out when a majority or a required percentage of participants agree to sign it [140].
2. Enabling automated transactions triggered by some specific events. This functionality can manifest itself in multiple ways, for instance, transactions automatically sent over at fixed time intervals (e.g., real-time requirements) or transactions sent in response to other transactions (e.g., feedback loop control cases). This facilitates request-response style transactions for decentralized data access within a blockchain-based system.
3. Allowing storage space for the application-specific information, such as membership records, lists, or Boolean states.

With well-written and secure smart contracts, many applications offer various functionalities, utilities, and algorithmic processing in blockchain networks. For example, Hawk is a smart contract-based platform designed to provide anonymous transaction services [141]. In general, the smart contract can provide IIoT applications with many advantages, including autonomy, trust, traceability, safety, efficiency, auditability, and accuracy. The deployed smart contracts are typically stored within the blockchain, rendering them available to all network participants. However, security lapses may occur if a participant exploits any bugs or loopholes in a deployed contract. For example, in June 2016, the DAO (Decentralized Autonomou Organizations) attacks in Ethereum networks resulted in the attacker unlawfully siphoning off Ether worth 60 Million USD, with transactions that were valid according to the exploited smart contract [70]. Thus, when deploying and dispatching a smart contract, the following matters must be addressed: 1) bug-free code and 2) government regulations and taxation.

V. CHALLENGES AND SOLUTIONS OF INTEGRATION

Both industrial IIoT and blockchain technology are still in their infancy stage, and many technical issues and challenges will arise upon integration. From future perspectives, blockchain will have a significant impact on the next generation of industrial IoT, although it still requires many efforts to standardize the architectures and policies for both blockchain and industrial IoT.

A. Challenges

Numerous challenges from different perspectives exist, such as technical challenges and social issues. In this subsection, we highlight the critical technical challenges produced by the integration process.

1) Technical Challenges: Most current blockchain prototypes are designed to run on P2P homogeneous networks. However, the unique characteristics of industrial IIoT (e.g., limited resources on end devices as compared with the high-performance servers or computing devices) prevent directly deploying blockchain into industrial IIoT. Several key challenges must be overcome when integrating blockchain into IIoT applications.

a) Computation: It is generally unaffordable to perform blockchain operations, with respect to higher computation and throughput requirements, on lightweight IIoT devices. However, some sophisticated cryptographic algorithms are used for privacy-preserving, such as Zero-Knowledge Proof [142] and Attribute-Based Encryption (ABE) [79], are still too heavy for these industrial IIoT devices. A full node in a blockchain-enabled IIoT (e.g., the gateways) should have the ability to verify and search for every block and transaction, which can also be a heavy task for resource-limited IIoT devices, even for gateways. Due to the limitations on computation and bandwidth, typically, PoW-like consensus protocols are not practical to deploy on lightweight IIoT devices. For instance, when running a typical consensus node, such as PoW in Bitcoin, on a modern Graphics Processing Unit (GPU), it can achieve about 10^7 hashes per second [143], and it is still very challenging to find a possible solution within 10 mins. However, even a powerful IIoT device (e.g., Raspberry Pi 3 [144]) can achieve only about 10^4 hashes per second [145]. Traditional IIoT devices are therefore unable to contribute adequate computational resources and afford these PoW tasks. This kind of situation can be found in the forms of PoW-like consensus protocols.

b) Storage: The massive storage required by blockchain nodes can be prohibitive for most IIoT devices. The participants of a blockchain with a small storage capacity will be in trouble, as blockchain is a shared data replication system, and storing all the
data blocks is necessary. Each node is required to have an exact copy of the data, which will definitely increase the storage costs on devices. However, without these massive data of blockchain, the IIoT devices will have difficulty or find it impossible to verify the validity of transactions generated by other peers. Also, to generate new transactions, a transaction sender requires the historical data (e.g., the balance and transaction index of previous transactions), which in turn requires that the IIoT devices know the current blockchain status. In this case, the IIoT devices have two options: either trust itself by adding extra storage or trust the remote servers. Also, the second option imposes extra communication overhead between the IIoT devices and the trusted servers.

To better understand this challenge, we provide a numerical comparison, regarding the storage issue, between Bitcoin and a medium-size Industrial IoT (IIoT) system. In Bitcoin, the block size is currently limited at 1MB. The average size of a Bitcoin transaction, in one week of February 2019, is around 500 Bytes [146]. Considering that the average number of transactions per block is 2000, and a Bitcoin block is generated by the miner approximately every 10 minutes, then every second, 3.33 transactions are generated within the Bitcoin network; thus, the average data volume is 1.67KB per second, which is pretty mild. We evaluate an industrial plant which has many wireless sensor and actuator networks (WSANs) deployed. We choose a medium-size system to estimate the average data volume, which consists of 50 WSANs, each having 100 nodes. We assume the average device sampling period is 1 second and the average message size is 100 bytes. This leads to an average data volume of 500KB per second. Assuming the block size is limited at 1 MB, then from the above comparison, we observe that in one week, the average block volume generated from the Bitcoin network is about 1 GB, while the average block volume generated from the medium-size IIoT system is 302.4 GB; this is a huge amount of data that certainly cannot be stored in local IoT networks. It is worth noting that the required data for immutable and verifiable services are application-dependent. Typically, these data will be stored for at least one year in the industrial case.

c) Communication: Blockchain leverages and runs on P2P networks as its underlying communication infrastructure, which requires participating nodes to frequently perform data transmissions and data exchanges. These nodes are required to keep exchanging the data in order to maintain the consistency of records (e.g., the latest transactions and blocks on its blockchain copy). However, in most industrial use cases, wireless communication technologies have already been widely used to connect IIoT end devices. The wireless connection may suffer from more challenges (e.g., shadowing, fading and interference, unreliability) than the wired connection [147]. The capacity and efficiency of current wireless communication technologies are far lower than the requirements of blockchain. For example, in a practical industrial use case, it typically adopts Bluetooth of IEEE 802.15.1, ZigBee of IEEE 802.15.4, Ultra-wideband (UWB) of IEEE 802.15.3, and Wi-Fi of IEEE 802.11 ab/g. These communication technologies, however, are far from fulfilling the general requirements of communication in a blockchain network. For instance, Bluetooth can provide a data rate of 250 kbps; UWB can provide a data rate of 110Mbps; Wi-Fi can provide a data rate of 54 Mbps [148]; and the new NB-IoT [149] can only provide a signal level of around 100 kbps [150]. These wireless technologies are a long way from fulfilling the requirements of general P2P communication.

d) Energy: Blockchain networks typically require more powerful devices for information processing and transaction verification. The energy consumption and maintenance costs of these devices are huge. However, in IIoT, many devices are designed to operate for long periods without directly connecting to power outlets, which means they are typically powered by batteries. For example, an IIoT device is designed to consume 0.3mWh per day, and it can operate for at least 5 years using a CR2032 battery with a capacity of 600mWh [150]. In a practical industrial use case, these IIoT devices may adopt more energy-saving approaches, such as sleep mode when idle [151] or high-efficiency communication technologies (e.g., NB-IoT [150]). However, the computation and communication in blockchain operations are typically energy-inefficient, which requires lots of energy to properly support the system function. When implementing blockchain operations into a typical IIoT device, the energy powered by a battery is used up very quickly, and correspondingly these devices will be offline. Taking the SHA-256 operation and ZigBee protocol as an example, the average energy consumption of a SHA-256 operation requires around 90 nJ/B [152], and the normalized communication energy cost for ZigBee protocol is around 300 mJ/Mb [148]; both are the basic operations (if integrating blockchain into IIoT). If the energy budget of an IIoT device is 0.3 mWh per day, as stated before, then it can only support about 0.5 MB data processing and transmission using the ZigBee protocol; this amount is far from fulfilling the requirements of blockchain [153].

e) Latency and capacity: Typically, when blockchain builds upon a consensus protocol to construct blocks (e.g., approval for transactions) and appends these blocks onto the chain, there is a waiting time requirement (e.g., 10-60 minutes or even longer for Bitcoin) to get approval and finalize the transactions among all participating nodes. This is typically for PoX-based consensus protocols. The high latency in these consensus protocols aims to ensure consistency in decentralized blockchain networks. This kind of latency is not acceptable for most mission-critical industrial IoT applications. The long block confirmation time (e.g., 10 mins) is also unacceptable for these time-sensitive industrial IoT applications (e.g., real-time applications). The long finality time (e.g., a time interval between transaction generation and completion) is a big challenge in making blockchain fit for smart factories [154]. BFT consensus protocol may provide a solution. However, it is subject to scalability issues, see Section IV-D.

f) Mobility and partition of IIoT: A typical industrial IoT network consists of two modes: 1) the stable network mode among the fixed infrastructures (e.g., base stations); 2) the ad-hoc mode, where the network does not have a pre-existing infrastructure and each node forwards data to its neighboring nodes [155]. Generally, the mobility on IIoT devices can heavily degrade the performance and efficiency of blockchain protocols, which require dynamic adjustments and configurations in communication. In other words, to maintain successful communication, the mobility in a wireless network can lead to an increase of control messages and signaling [156]. In addition, wireless ad-hoc networks usually partition the overall network into multiple disconnected sections when mobile nodes switch with diverse patterns [157]. Both consistency and synchronization among these mobile nodes are the biggest challenges in resource-constrained IIoT devices.

g) Timestamping Authority: A basic blockchain transaction automatically includes a time-stamp field, which indicates when this block or transaction is created. An important issue with the existing blockchain infrastructures, which must be resolved in order to fit industrial IoT applications, is the lack of obligation with a time-stamping authority or the authority for time assessment. An accurate assessment of time is crucial for any industrial IoT, but blockchain lags behind in this case since it needs time to get consensus among the participating nodes [158]. This kind of latency is very common before the block gets its finality due to the consensus procedure. Also, considering the transmission delay, this issue is more practical. Besides the above-mentioned challenges, there are other traditional
technical-related challenges for both IIoT and blockchain, such as security vulnerability, privacy leakage, etc.

2) Standardization for Blockchain-enabled IIoT: By offering new features on decentralization and immutability, blockchain technologies have greatly revolutionized the industries conceptually. The advent of this innovative technology has great potential to reshape the whole current IIoT market. The integration of blockchain into industrial applications is still in its infant stage, and there are many technical challenges and issues urgently required to be resolved before successfully integrating them together. One of the most fundamental things needed for integrating these two technologies together is to set up the rules and policy for both blockchain and industrial IoT [159]. Although the integration of blockchain and IIoT can bring numerous advantages, this integration has developed without any standards and is currently limited to only a few service providers. Vendor-specific blockchain technological advances are coming from distinct research organizations, banks, and factories, each with its own independent policies and architectures. To adopt this integration in different industrial IoT applications, the integration process must be standardized and must follow some specific patterns for future compatibility [160]. Until now, both blockchain and IIoT are the vendor and use case-specific in independent production systems. Each service provider mainly designs and offers its own solutions for its specific applications rather than providing a generic and standard scheme that can be applied to diverse use cases.

Typically, the lack of standards on integration restricts the potential collaborations between different services providers; it creates difficulty for customers in changing and choosing the providers, as each provider may have distinct rules [161]. In addition, the non-standard heterogeneous communication protocol between various blockchain platforms and industrial IoT applications remains a critical issue for the entire industry landscape. To create a successful industrial IoT environment with blockchain technology, the independent approaches must be replaced by an open, transparent, and standardized policy. In order to obtain proper services from blockchain technology, service-level agreements between various industry sectors are also necessary. If failing to finalize the standardization, technologies will grow independently, causing serious trouble in future incompatibility.

To reach an agreement on the integration of blockchain and IIoT, technical details (e.g., network setting, blockchain deployment, IIoT device integration and configuration, services payment schemes) should be carefully considered. Federation of service providers can actively take this role to standardize this integration process. A number of international organizations, such as ISO, ISTIC Europe, and IEEE, have contributed to standardization efforts in building general functional architectures for blockchain-enabled IIoT platforms; however, these efforts are still in their initial stages [162].

B. Potential Solutions

This section discusses the potential solutions to the above-mentioned challenges and issues. Although these solutions cannot handle all challenges and issues, they can be used to mitigate many challenges when integrating blockchain into IIoT platforms.

1) Transaction Format: The format of transactions matters when dealing with the higher latency issues in traditional blockchain protocols. Different from transactions in crypto-currencies, transactions in industrial applications need to support user-defined data structure [163]. Although several practical examples have been built on Ethereum [164] [165] for IoT applications, those examples are still far from meeting industrial requirements (e.g., timing constraints). In Section IV-C2, we proposed a transaction format for industrial smart metering applications. Note that for different industrial applications, the transaction format may have different fields. However, the transaction in IIoT has at least one data field indicating the data to be transferred. The data field typically has varying lengths, and a sender may be required to pay a much higher transaction fee for a longer data field. However, considering the network traffic and practical situations (e.g., network communication protocols), the length of the data field cannot be enlarged unlimitedly.

Transaction size, particularly under IIoT networks with unreliable wireless channels, can significantly affect the confirmation delay of transactions. Typically, a small transaction can have a high transmission success rate and low transmission delay. Many industrial applications are still using the User Datagram Protocol (UDP, a lightweight protocol) as their communication protocol [166]. However, the UDP protocol typically only provides a basic transmission, for example, it does not provide an error-correction mechanism, due to extra overhead. For lightweight industrial applications, special care is required. For example, to minimize the network frame fragmentation and improve the transmission success rate, it is better to maintain the transaction size as less than the payloads of network protocols (e.g., UDP and IP). These smaller transaction sizes have a higher probability of being mined by the participating nodes (also commonly called “miner”).

In addition, the delay can be further optimized with powerful agents or devices (e.g., edge gateways) that can wirelessly establish the connection with industrial IoT devices and then wirely connect itself to the miner. The agents can equally broadcast transactions of different sizes to the miners.

2) Incentive Mechanisms: The incentive is the most common practice for motivating the participating nodes to actively participate in a predefined consensus protocol and make them “work hard”. Typically, a transaction is associated with a transaction fee in blockchain, which is an essential mechanism for balancing transaction costs and adjusting the resource consumption of blockchain. For example, the transaction fees can be used to indicate the complexity or urgency of transactions [115]. Typically, the transactions consuming more resources (e.g., computation, communication, and bandwidth) incur higher transaction fees. In addition, the transaction fees can also provide a mechanism to fairly reallocate network resources, especially for capacity-limited public blockchains. In the case where a large number of transactions are generated at a moment, these transactions may suffer from a much longer transaction confirmation time. To resolve this conflicting competition, the sender of the transactions can pay more transaction fees to incentivize the miner to give priority and better services to those transactions (e.g., shorter confirmation time).

In an industrial application, the transaction fees may not be a monetary-related cryptocurrency, and they can be in the form of tokens. A token system in blockchain can be used as a reliable reputation or trust system to incentivize the miners [167]. For industrial IoT networks, the incentive mechanism is attractive and non-negligible. It can, to a certain degree, increase the cost of attacks as compared to traditional IIoT attacks (e.g., forged messages and DoS attacks), which can further prevent malicious behaviors [115].

Due to limited resources and poor wireless links, industrial IoT devices may not be able to mine blocks and earn a token for the transaction fees. However, these devices can “sell” their services for the tokens (e.g., the rechargeable energy or the chance to be charged). For instance, a service user, such as the IIoT administrator or cluster header, can recharge the energy for IIoT devices based on the tokens a device obtained. The IIoT devices are expected to actively participate in blockchain consensus procedure and obey benign behavioral patterns. Combining with the deployment of smart contracts, the IIoT devices can purchase resources (e.g., power or data pack) for more rewards, which can motivate IIoT devices to earn more
tokens.

3) Smart Contracts: A smart contract is a piece of “secured execution of code” [138] [168] that executes without any assistance from a trusted third party. Once the specified conditions are fulfilled and triggered, the smart contract self-executes the corresponding contractual clauses. Also, the smart contract can act as a real-time auditor, since all actions are required to be recorded and verified as a transaction in a decentralized blockchain. These transactions are typically trackable and undeniable, which enhances machine-execution security [169]. It also can translate various assets (e.g., IoT devices and digital assets) into virtual identities in blockchain and enable them to interact with other assets. Typically, the script of a smart contract is stored in the blockchain and can be identified by its unique address. In general, there are two options for executing a smart contract: a) the receiver validates the transactions with smart contract address; or b) the internal execution of code [115]. Using blockchain, all execution records can be tracked. The smart contract is executed independently and automatically on every node of a blockchain network. These features of smart contracts (e.g., automatic execution with defined rules in a decentralized manner) have the potential to improve the efficiency and security of an IoT application, as they do not involve human interference. For example, with a smart contract, blockchain has the potential to replace the Intelligent Transportation Structure (ITS) and realize a reliable firmware update on IoT devices [170].

4) Off-Chain Storage: Due to the feature of decentralization, the block data should be duplicated among all participating nodes; this poses a great challenge on storage for resource-constrained devices. Off-chain storage is a potential solution for mitigating this challenge. Instead of storing all data on-chain, the actual industrial data (e.g., measurement records) can be securely stored separately at another place; a pointer then points to the index in the blockchain. There are several works on off-chain storage solutions in the literature. For example, in [81], two different types of transactions are introduced, namely, the transaction for access control management and the transaction for data storage and retrieval. An off-chain key-value store is an implementation of Kademillia [171], a distributed hash table (DHT). The DHT is managed and maintained by all network nodes that are independent of the blockchain process. Data can be randomized across nodes and replicated to ensure availability. In practice, the cloud in traditional cloud-based IoT platforms can be an ideal off-chain storage site, which provides virtually unlimited storage capacity [30].

5) Other Data Structure: In addition to traditional blockchain structures, there are many different blockchain-like structures that can provide decentralization and immutability services, such as DAG (Directed Acyclic Graph) [172], BlockDAG [173], and GHOST [174]. These structures typically focus on achieving scalability and higher throughput in IoT application domains, which can also be applied to industrial use cases as well with some modifications.

Different from the main chain solutions in traditional blockchains, DAG solutions reorganize the blocks, rather than a single sequence of blocks, in a finite, directed graph with no directed cycles. One example is BlockDAG, based on a DAG solution. In BlockDAG, the vertices represent the blocks, and the edges represent the links between multiple previously published blocks and the current block. BlockDAG does not aim to eliminate PoW mining scheme or transaction fees, but instead leverages the novel structural properties of DAG to reduce the higher orphan rates in blockchain systems. Another practical DAG example is the Tangle [85], which is built directly on transactions instead of blocks (a set of transactions). In Tangle, a transaction must approve two previous transactions before appending its “block”. Unlike the single-copy in chain structure, Tangle does not drop any transactions, not even the conflicting transactions, and instead keeps all transactions in branches of the DAG. Typically, a DAG structure can achieve better capacity and scalability; however, it has some internal drawbacks, for example, weak consistency. GHOST (Greedy Heaviest-Observed Sub-Time), another blockchain-like structure, organizes blocks in a tree structure [175] [174]. It takes the path from the genesis block (the root of the tree) to the heaviest sub-tree (having the maximum number of blocks), which indicates that it contains the heaviest computation quantity as the publicly accepted main chain.

This section discusses several potential solutions for solving the challenges on integrating blockchain into IIoT platforms. Besides the structural aspects, there are many other optimized schemes (such as consensus protocol) to further mitigate these challenges, e.g., applying communication efficient BFT protocols [84] to improve the response time.

VI. Practical Industrial Applications

The integration of blockchain and IIoT has prompted the appearance of a new set of smart services and applications that bring substantial benefits to industrial processes. There is a wide range of research in the integration of blockchain and IIoT in many smart industries. This section provides a brief summary of several key applications across different scenarios, including Industry 4.0 in general, smart manufacturing, smart grid, smart energy, supply chain, and the food industry. The selected industrial use cases are representative, considering current industry domains. We will focus on the opportunities brought by the integration of blockchain and these industrial scenarios. In addition to industrial IoT applications, it can also enhance and extend to general IoT applications’ scenarios, such as smart healthcare, smart city, smart transportation, smart agriculture, as well as service-related applications (e.g., smart education, smart cloud services). This section mainly focuses on the industrial use cases.

A. Industry 4.0

With the advance of the automation industry, the complete automation of industry and business processes has become a reality. An abundance of technological advances and their integration into the industry has led to an emergence of new approaches to production, known as Industry 4.0, incorporating the prowess of various technologies, such as IoT, blockchain and cyber-physical systems (CPS) [176]. Industry 4.0 is expected to offer existing industrial systems with promising transformation. It has been considered as a key enabler for the next generation of advanced industrial automation systems [177].

In the current competitive market, companies aim to earn business advantages at any cost via agents (can be viewed as the participating nodes in a blockchain). To minimize the communication costs and potential risks among them, a possible solution is to have these agents communicate directly. However, this raises the question of trustworthiness between the participating agents. The use of decentralized systems, such as blockchain, can be a promising solution for efficient and secure communication between autonomous agents [178] [179]. The use of blockchain in these business models can provide trustworthy and immutable services among the involved parties. In general, deploying blockchain into these business models can provide several advantages: 1) build trust between parties and devices, reducing the risk of collision and tampering; 2) reduce cost on the overhead associated with middlemen and intermediaries; 3) reduce latency of the settlement time from days to near-instantaneous.

In Industry 4.0, the real-time QoS monitoring is an essential part of any modern business process. Unlike some crypto-currency
With increasing demands on energy usage to support industrial and manufacturing operations, smart energy and its management system continue playing an integral part in most industry ecosystems. The emergence of distributed renewable energy resources is reshaping the role of consumers from pure consumers to prosumers, who can generate energy (e.g., from renewable energy resources) in addition to consuming it [183]. The energy in transition can be in any form, such as electricity, gas, and heating grids. Energy prosumers who have extra energy can sell them to other consumers who need them. The energy trading process between the prosumers and the consumers (e.g., both as the peer nodes) can be in a form of P2P energy trading. However, it is challenging to assure secured and trusted energy trading between multiple trading parties in a distributed and decentralized manner.

The centralized energy management systems (EMS) appear to not work efficiently with a large number of prosumers; thus, a decentralized architecture based on blockchain technology is necessary to achieve a high quality of services for various energy entities in a decentralized manner [184]. The goals of the blockchain-enabled smart grids are not only to create a trusted, reliable, and efficient smart energy network, but also to improve security and privacy among energy exchange and transmission. The appearance of blockchain technologies brings great opportunities for ensuring secured P2P energy trading; moreover, some recent studies proposed to use of blockchain to handle the challenges in the EMS. For example, Xu et al. [185] proposed a blockchain-based crowdsourced energy system (CES), facilitating a P2P energy trading at the distribution level, where ubiquitous distribution-level asset owners can trade with each other without the help of trusted authorities. The trading platform is implemented in the IBM Hyperledger Fabric network, which is deployed in multiple clouds to offer the required blockchain services. In their proposed platform, smart contracts are used to run the pricing mechanism as well as control energy trading transactions and crowdsources. In addition, it also proposed an intelligent energy-aware resource management system within cloud data centers using blockchain technologies. One of the goals of their platform is to minimize the cost spent in cloud data centers and reduce the cost of energy consumption from the traditional power grid, the request scheduling cost, and the request migration cost in data centers. In general, blockchain technology has great potential to improve both the security and privacy of energy exchange and transmission, while the cloud offers unlimited storage and powerful management services, and supports blockchain to achieve decentralized energy operations.

D. Supply Chain

Typically, an industrial product consists of collaborative work from multiple suppliers that are from different manufacturing sectors across countries. However, some forged (e.g., low-quality or reused) parts may seep into the supply chain system. Applying anti-fraud technologies in every part of products is quite expensive. The integration of blockchain and IIoT can resolve this issue. Normally, each part associates with a unique ID when it is created. An immutable timestamp is then attached with this ID. The identification information of each part can be recorded into a blockchain as tamper-resistant proof. For example, Konstantinidis et al. [186] claimed that the ownership of the part in a product can be authenticated and recorded through a blockchain-based system; Kim et al. [187] presented a traceability ontology, with the integration of blockchain and IoT technologies based on Ethereum blockchain platform, to provide tamper-proof evidence for products.

Blockchain-enabled applications can also be used to reduce the costs of after-sale services in supply chain management. For instance, Tapscott et al. [188] showed some use cases of motor insurance, in which the settlement of claims can be automatically executed via smart contracts based on a blockchain setting, thereby improving efficiency and reducing claim-processing time in traditional cases.
The blockchain can enhance the visibility of the life cycle of products, especially in the food industry. The traceability of food (including its ingredients) is a necessity to ensure food safety. However, it is a challenging task for the incumbent IoT to guarantee food traceability throughout the entire food supply chain [189]. Typically, traceability requires digitizing all information, including the raw materials, from the sources to every sector of food manufacturing. Blockchain can ensure this kind of traceability and provenance of food industry data. There are several case studies in this area. For example, Tian et al. [190] proposed to use RFID and blockchain to create a supply chain platform from agriculture to food production. This system has demonstrated a guarantee of the traceability and safety of food supply chain data. Also, these data are immutable when empowered by blockchain. In general, regarding the type of blockchain, the food industry can adopt a consortium blockchain to enable many different industrial sectors to work together.

VII. DISCUSSION

This section provides some discussion on existing Blockchain-as-a-Service (BaaS) platforms, on how to choose the right blockchain for a specific IIoT project, or alternatively, to answer the question: do we really need a blockchain?

A. Blockchain Storage and BaaS Platform

The integration of blockchain and industrial IoT provides an unprecedented architecture for enabling smart services across industrial IoT domains. As described in Section II-C, current IIoT applications largely rely on the cloud to provide blockchain storage. This subsection discusses the current potential decentralized cloud-based blockchain storage and popular BaaS platforms.

1) Decentralized Cloud Blockchain Storage: Data storage in traditional cloud-based IIoT platforms has largely relied on cloud computing which is a fully centralized structure. However, the centralized storage architecture usually faces several serious challenges, such as the lack of user control over the IIoT data, security and privacy concerns, as well as a high fee from service providers. On the other hand, it appears that the conventional blockchain system is very expensive to store large amounts of IIoT data on-chain, especially for resource-constrained IIoT devices. To overcome such obstacles, decentralized storage that relies on both cloud and blockchain technologies is a promising solution, offering highly flexible, secure, trustful, and super cheap storage services for IIoT applications [191] [192].

With this regard, we survey several popular decentralized storage platforms, as shown in Table IV. Table IV provides the key technologies and open sources of their software. We refer interested readers to read the corresponding white papers. With decentralized storage platforms, cloud-based IIoT applications do not need to rely on a centralized service provider, which allows users to store their IIoT data to a set of distributed storage nodes (e.g., P2P nodes). In fact, many of these platforms provide efficient IoT solutions for industrial use cases. For example, compared with traditional centralized storage solutions, the decentralized IPFS [193] and StorJ [194] storage platforms demonstrate efficiency and robustness when implementing into IIoT systems, e.g., in terms of low access latency and improved security levels.

2) BaaS Platform: Recently, some well-known cloud providers have launched initiatives to integrate the decentralized storage services for large-scale, cloud-based blockchain platforms (e.g., IPFS storage on Amazon and Microsoft Azure clouds [192] [213], to achieve secure and efficient data storage and management. Blockchain can be regarded as a Blockchain-as-a-Service (BaaS) on these platforms, leveraging the cloud to offer full IT services, which helps customers to build, create, validate, and deploy blockchain for their cloud-based IIoT applications. Typically, BaaS platforms are capable...
of providing some foundational infrastructures and technical supports to ensure that the target cloud-based IIoT systems can achieve robust and efficient operations.

Table V shows several potential BaaS platforms for general cloud-based IoT applications. These BaaS providers enable customers to quickly develop and deploy the required services without worrying about the underlying infrastructure installation and system investments, which potentially accelerates the deployments in practical use cases. Most of their source codes for BaaS examples and templates for deployment are available and accessible on code-sharing platforms, for example, Github [214]. Many active research projects have deployed their BaaS platforms for developing various IoT applications. For example, IBM cloud-based IoT platform integrates with IBM BaaS services to manage vehicle sensor data and ensure security and privacy during the data sharing process within vehicular networks [215].

Although the development and deployment of these BaaS platforms are still in progress, the success of such initial projects on BaaS platforms is expected to open up many new opportunities for future cloud-based IIoT deployments as well as re-shape future industry markets.

### B. Distributed Decentralized Systems

Blockchain as a distributed and decentralized ledger offers some unique opportunities when it is integrated into IIoT applications. It enriches the pure competing databases or the traditional distributed systems, which explores a new methodology on the system design. With these unique features, integrating a blockchain on existing platforms, however, comes with a significantly high overhead (e.g., replicating all data records and the corresponding operations at every participating node of the system, even when they are only occasionally participating in operations). To better provide services for various applications, when trying to integrate blockchain into IIoT applications, it is better to ask several critical questions to establish whether the blockchain technology would be a good fit for a specific project.

These questions primarily consider several aspects, such as governance, operations provenance, and attacks to prevent [216]. For example, if the designed system requires neither shared governance nor shared operations, then consensus protocols associated with blockchain are likely unnecessary overhead for that system. On the other hand, if the designed system requires both shared governance and shared operations, then blockchain may be a necessary and
feasible structure to adopt. Other parameters must also be considered to decide whether blockchain really fits a project.

According to different design principles and usages, most existing distributed or decentralized systems can be roughly classified into one of five cases: replicated database, replicated monitored database, replicated monitored ledger, replicated database with consensus, and blockchain. Essentially, blockchain is a type of distributed and decentralized database that fits within the broader family of distributed systems. Table VI shows a comparison on different categories, which helps the readers decide if the blockchain suits a specific project. The table focuses on four distinct considerations to classify them.

The first consideration regards what is the operation model? This consideration is used to decide the owner of the database (or the question of who has the right to operate on the database). This question has two distinct options: singular (e.g., only a single entity (the owner) can operate on the database) or shared (e.g., multiple entities can collaboratively operate on the shared database). In shared mode, the system would require a consensus protocol to allow for shared governance, so that the database can maintain consistency among multiple entities. The second consideration regards where is the root of trust? This question is trying to answer who will be responsible for honesty among the involved entities, to ensure the system works correctly and securely. This question also has two options: maintainer or system. Maintainer means that the trust is rooted in the maintainer (e.g., if the database is using the Microsoft Azure cloud storage [see Table V], then it requires that the participants must trust Azure cloud). System means that the trust is rooted in the design of the system itself. This option is possible only if the system reserves sufficient provenance for it, and some mechanism exists (e.g., consensus protocol) to guarantee the system functions as intended. Otherwise, the system may be compromised.

The third consideration regards what is auditable? This question typically has three options: nothing, current state, or provenance. Nothing means nothing is required to be auditable, which is the worst case. Current state means that system can resort to an authenticated data structure [217] to ensure that its current state can be audited. If the state also contains a history of the system (e.g., via a ledger), then the use of an authenticated data structure allows for the provenance of the system to also be audited. This is the option provenance. In either case of current state or provenance, the database is required to be monitored to guarantee that the systems are never entering an invalid state, even temporarily. The fourth consideration regards what is resilient against? This question is trying to answer what kinds of malicious behaviors can be avoided for consistency of the database. Typically, it has three considerable resilient properties: data loss (e.g., is it resilient to an accidental data loss?), detect (e.g., is it possible to detect that data has been maliciously altered?), and prevent (e.g., is it possible to prevent malicious updates?) [216].

C. Integration Approaches and Blockchain Selection

It is publicly agreed upon that integrating blockchain technologies into IIoT will offer many advantages to IIoT applications. However, some challenges and disagreements still remain; for example, one of the key issues is where blockchain should be hosted [218]. Both blockchain and IIoT have their unique infrastructures and requirements. For instance, due to the resource constraints (e.g., on computational resources and bandwidth) of current IIoT platforms, it is inadvisable and impractical to directly integrate blockchain into IIoT applications. In addition, regarding its computational resources and latency, the service-level platforms (e.g., a cloud and a fog on edge networks) can be a potential integration platform. Compared with the cloud, the fog may have limited resources; however, it typically exhibits much lower latency. While cloud-based platforms can scale-out which can serve as centralized management, and also can overcome resource constraints at the cost of significant latency [219]. Based on these characteristics and limitations of industrial devices, many promising models have been proposed for the integration of blockchain and IIoT. Generally, these models can be classified as three ones [220].

a) IIoT-IIoT: This model typically focuses on functionalities of the IIoT side, and blockchain is typically as an immutable database. The duty of blockchain is limited, in which it only requires occasionally to access or interact with the blockchain. In this model, only part of IIoT data is added to the blockchain, whereas the interactions among IIoTs typically happen without blockchain. This model typically requires some proxies (e.g., gateways) to connect the IIoT part with blockchain. This model is more practical for some mission-critical applications (e.g., real-time applications, which require reliable and low-latency interactions among IIoTs).

b) IIoT-Blockchain: This model involves a huge communication and interaction between IIoT and blockchain, which all data records and interactions are required to record in blockchain in an immutable and traceable manner. This model should provide an interface for direct communication between IIoT and blockchain, and it has a high requirement on bandwidth to transfer data and perform consensus procedures. This model is more practical in financial-related scenarios, where data records are more valuable resources. However, recording all the interactions into blockchain would require a huge bandwidth and computational abilities.

c) Hybrid Approach: This model only records part of the interactions into the blockchain, while other interactions are directly shared between IIoT devices without being included in the blockchain. By doing so, it does not require heavy communication between IIoT and blockchain (e.g., only recording some critical data into the blockchain). However, one issue is what kinds of interactions should go through the blockchain, and how to provide a way to make this decision in a run-time manner. This solution requires a careful system design to choose interactions (e.g., via labeling technologies). In particular, this model is a good candidate for leveraging the benefits of both blockchain and real-time IIoT interactions. Considering the current challenges of both IIoT and blockchain, this method may be a promising solution in the near future.

Table VII summarizes IIoT application requirements on the above integration approaches. Table VII, consisting of throughput, latency, security, and resource consumption, offers a comparable view on their strengths and weaknesses to help system designers choose the right one. In general, blockchain technology, with the help of consensus protocols and smart contracts, can serve as an enabler for IIoT to provide secure, reliable, and immutable data storage.

With the diversity of solutions on the integration of blockchain and IIoT (e.g., different types of IIoT devices and applications), it is critical that system designers select the appropriate and suitable solution based on their own restrictions and requirements. When integrating blockchain into IIoT, the IIoT platforms require blockchain to store and record their states, manage multiple “writers” (or data producers), and prevent the use of a trusted third party. Fig. 9 shows a simplified flowchart that can be used to help system designers to determine which kind of blockchain is suitable for IIoT applications [221] [47] [222].

VIII. RESEARCH TRENDS AND FUTURE DIRECTIONS

This section discusses some research trends and future directions for integrating blockchain into industrial IoT. Instead of discussing future directions in blockchain or IIoT separately (e.g., consensus protocols and smart contracts of a blockchain, or security and resource
TABLE VI. COMPARISONS OF DISTRIBUTED DATABASES

<table>
<thead>
<tr>
<th>Databases</th>
<th>Operation Model</th>
<th>Root of Trust</th>
<th>Auditability</th>
<th>Resilient Against</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicated Database</td>
<td>Singular</td>
<td>Maintainer</td>
<td>Nothing</td>
<td>Data Loss</td>
</tr>
<tr>
<td>Replicated Monitored Database</td>
<td>Maintainer</td>
<td>Current State</td>
<td>Data Loss</td>
<td>Detect</td>
</tr>
<tr>
<td>Replicated Monitored Ledger</td>
<td>System</td>
<td>Provenance</td>
<td>Data Loss</td>
<td>Detect</td>
</tr>
<tr>
<td>Replicated Database with Consensus</td>
<td>Shared</td>
<td>Maintainer</td>
<td>Current State</td>
<td>Data Loss</td>
</tr>
<tr>
<td>Blockchain</td>
<td>System</td>
<td>Provenance</td>
<td>Data Loss</td>
<td>Prevent</td>
</tr>
</tbody>
</table>

TABLE VII. COMPARISON OF INTEGRATION APPROACHES

<table>
<thead>
<tr>
<th>Throughput</th>
<th>Latency</th>
<th># of Untrusted Writers</th>
<th>Data Storage</th>
<th>Consensus Mechanism</th>
<th>Security</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Database</td>
<td>Very High</td>
<td>Fast</td>
<td>Cloud</td>
<td>None</td>
<td>High</td>
</tr>
<tr>
<td>IIoT-IIoT</td>
<td>Low</td>
<td>Fast</td>
<td>High</td>
<td>BC/IoT Devices/Fog</td>
<td>PoW, PoS</td>
</tr>
<tr>
<td>IIoT-Blockchain</td>
<td>High</td>
<td>Slow</td>
<td>Low</td>
<td>Blockchain (BC)</td>
<td>PoW, PoS, BFT</td>
</tr>
<tr>
<td>Hybrid</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>BC/IoT Devices/Fog</td>
<td>BFT</td>
</tr>
</tbody>
</table>

Constraints of IIoT, we discuss future directions on the integration of these two technologies.

A. Optimization on Performance

Considering the distinctive features of the integration of blockchain and IIoT, most current integrated solutions do not provide a decent performance (e.g., low throughput and high latency), which is not preferred by their host applications. In an IIoT network, a large number of lightweight devices are required to simultaneously communicate with each other, necessitating a network with a high throughput. This is an easy task with the help of a centralized controller in the IIoT platform (e.g., the Supervisory Control and Data Acquisition (SCADA) system [223]). However, blockchain technology requires some basic configurations on distribution and decentralization. In existing implementations and deployments, increasing throughput typically decreases scalability, which is not desirable in most industrial designs. Typically, system designers must consider scalability when their businesses/systems become large-scale. In a practical IIoT network, various devices need to communicate with each other in an almost real-time manner. The required latency in an industrial use case should be very low to fulfill the timing requirements. However, reducing latency typically compromises scalability, which is not acceptable in large-scale applications. Also, most IIoT devices are resource-constrained, while most current blockchain protocols require some sophisticated cryptographic primitives, which places a huge burden on these resource-constrained devices.

To successfully apply blockchain into IIoT platforms, practical solutions are required to improve the performance without sacrificing the scalability, so that the integration can be scalable to large networks and yield high throughput and low latency even with resource-constrained IIoT devices. Thus, some new design frameworks are urgently required to deploy and optimize the performance in large-scale scenarios. To achieve the required performance, it also requires balancing the trade-off between the affected factors, such as scalability vs. security. In addition, many new blockchain-similar structures, such as DAG, are introduced to tackle the performance limitations on throughput and latency. However, those structures are also subject to their own weaknesses, such as no fixed internal structures or difficulty in verifying a specific transaction.

B. Scalability

The key features of the blockchain (e.g., decentralization and immutability) require that every full node store a full copy of blockchain; however, this comes at a cost of scalability. The scalability issue in blockchain limits its wide usage in large-scale IIoT networks. Typically, the scalability can be evaluated by the throughput (e.g., measured by the number of processed transactions per second) against the number of IIoT nodes and the number of concurrent workloads [224] [225]. In the current design, many blockchain systems are still suffering from poor throughput. Scaling blockchain has become an active research area [226], for example, via increased
block size [61] or sharding techniques [62]. Blockchain scalability issues are still an open research area, and many different initiatives and efforts in recent research are aimed at improving blockchain scalability, from side chains to sharding techniques [62].

Adding to the ubiquity of IIoT networks, certain industrial devices may be able to travel long distances, such as those installed on aircraft, international trains, and ships [227]. The integrity of the data generated by these mobile IIoT devices is equally important to those generated by static IIoT devices. However, the data generated by mobile devices may not have the ability to be recorded into blockchain in a short time (e.g., due to the loss of internet connection). The side chain technology [228] [229] offers a potential solution for transferring assets between multiple blockchains. With the help of side-chain technology, the data can be transferred among different chains in a decentralized manner. On the other hand, sharding [230] is a novel mechanism for enabling transactions to be processed in parallelization at a small scale. By paralleling, the block generation rate, and thus throughput, can be significantly improved. The data in a typical IIoT application may exhibit strong locality and heterogeneity, deeming it useful only to local regions, which provides great opportunities for developing sharding blockchain in IIoT domains.

Until now, scaling the blockchain remains a major challenge in their implementations in many industrial applications, due to their low performance and networking overhead. The issue surrounding low throughput is exacerbated in IIoT scenarios, where a much higher volume of data transactions occur simultaneously (data creation or transfer) and require a time-efficient manner to handle these data. One potential direction is the vertical scaling of blockchain as a decentralized database [54]. Horizontal scaling (e.g., sharding) also shows much promise in solving blockchain scalability issues, and the atomic commitment in inter-blockchain communication is another key research direction. Solving scalability in blockchain will serve as a huge advance toward creating a practical decentralized infrastructure for IIoT applications.

C. Security and Privacy

Although introducing blockchain into IIoT can potentially improve the security of IIoT applications via some robust encryption primitives and digital signatures brought by blockchain, the security issues are still a major concern for this integration due to the vulnerabilities of IIoT systems and blockchain systems, which may be different. The inherent security features of blockchain cannot cover all vulnerabilities in IIoT applications. Typically, a weak link may prove to be an exportable loophole within smart contracts. For example, Atzei et al. [70] shows that Decentralized Autonomous Organization (DAO) attacks occurred by exploiting the shortcomings in smart contracts. One direction that research must take for the successful integration of blockchain and IIoT is to develop security standards for scripting smart contracts; these standards should be written in such a way that no loopholes exist that would compromise the security of IIoT networks.

On the other hand, there is a growing trend in deploying wireless networks into industrial environments, taking advantage of both the feasibility and scalability of wireless communication systems. The open wireless medium, however, also leads IIoT to suffer from the security vulnerabilities, such as eavesdropping, jamming, or replaying attacks [231] [232]. Also, due to resource constraints in traditional IIoT devices, conventional heavy-duty encryption algorithms may not be feasible to IIoT [233].

Similarly, blockchain technologies have some mechanisms for preserving a certain degree of privacy for transactions recorded in blockchain (e.g., via anonymous identity technology). However, these adopted privacy-preserving technologies in current blockchain systems are not robust enough. For example, the attackers can track the user’s IP address [234], and the privacy breach can occur by drawing interference based on a graph analysis of network nodes with which a user transacts [235] [236]. A better solution for preserving privacy in the blockchain would be in a form of decentralized record-keeping that is completely obfuscated and anonymous by design. Several techniques can be used to mitigate privacy issues in blockchain, such as ring signature [237] and address mixing [238] [239]; however, when applying these techniques directly to industrial domains, they are also subject to other critical issues (e.g., resource constraint issues in performing complex computations in IIoT devices).

The challenges in designing a blockchain-based industrial IIoT platform that maintains both accountability and privacy have inspired many solutions, yet remaining open to further research and development. Many solutions rely on implementing access policies either within the blockchain itself or through smart contracts of blockchain. However, the design of an efficient privacy-preserving scheme for blockchain is still an active and open question.

D. Editable Blockchain

The storage in industrial IIoT devices can be very limited compared with the explosively growing size of the blockchain, as all records on the blockchain should be kept in every node in a long term. Typically, the data volume generated by a general IIoT application is much larger than most cases of cryptocurrencies. Even in the case of Bitcoin, since its genesis block in 2009 [240], its total data size has grown to 340GB by April 2021 [237]. However, after a constant duration (e.g., one year in condition monitoring), the data volume generated by the IIoT applications may be unimportant or meaningless, and this data volume can be stored on backup storage. For example, in the food industry, the food record (including the raw materials and preparing processes) is meaningless after the food has been consumed by customers; to reduce the storage, such data can be deleted from the blockchain or stored on backup storage. In addition, the fraud actions and records on the IIoT blockchain raise the demand for editable blockchain technology without breaking the trust of stored data. It is desirable to design an editable blockchain, which enables deletion or modification on some blocks when satisfying some specific and well-defined conditions. Due to the contrary to the inherent feature of immutability in blockchain, the editability of the blockchain should guarantee secure conditions and records for any edit actions. Even if a blockchain is designed to be editable, however, it must still guarantee immutability and consistency. When designing and developing an editable blockchain, it needs to balance the trade-off between the security of the system (e.g., the choice of a hash function) and editability. In this kind of blockchain, we grant the right to correct the wrong records while still enforcing its features as a blockchain (e.g., the global consensus among all participating nodes).

E. Edge Computing

Most IIoT applications have strict computational and networking constraints (e.g., timing requirements), which may pose some issues when using blockchain-based decentralized architectures. Even if these IIoT devices are incorporated into a blockchain system where the devices do not have the capability to mine new blocks, these devices typically do not come with some storage requirements to host a complete copy of a blockchain. In addition, various IIoT devices typically suffer from limited interoperability and a lack of authentication and authorization standards to follow. As an extension to the cloud, edge computing (alternatively called fog computing) has emerged as a promising technology to empower blockchain-based
IloT platforms. Typically, edge servers are not as powerful as clouds; however, these edge servers are located at the edge of the network with close proximity to the IloT devices. This enables highly efficient IloT data computation with much lower transmission delay, which can provide an instant response (as opposed to cloud [241] [242]).

When integrating IloT devices into the blockchain system, the powerful gateways can be the consensus nodes in the blockchain. However, the issue remains that the degree of decentralization achieved is still limited. A potential research direction can be to extend blockchain to IloT edge, which can limit the computational and networking overhead of resource-constrained industrial devices. And with these edge devices, the consensus can be achieved via end-to-end communication over blockchain via computationally capable IloT gateways. However, this would enable the IloT devices and gateways to push transactions to the blockchain using lightweight clients, without creating centralized block validation pools. This also requires new design frameworks and real implementations in practical IloT applications in order to test their performance.

F. Standardization on Blockchain-based IloT

There is currently a lack of standards for establishing compatible architectures on the integration of blockchain into IloT. Without available standardization, it is difficult or impossible to achieve a service agreement on these integration processes. Moreover, each organization may develop incompatible standards among these partners. Although currently, no standards exist, many standardization efforts have been made among the participating members (such as ISO and IEEE). For example, ISO approved Australia’s proposed international blockchain standards in 2016, and the standard for blockchain and distributed ledger technology (ISO / TC 307) was released in 2019 [243] [244]. In addition, many initiatives on the development of blockchain-related standards are still in progress, covering major blockchain topics such as terminology, privacy, governance (AS ISO/IEC 37500), interoperability, security, and risks. In industry, the U.K. and Europe have developed many standards to support the scenarios in financial transactions and the role of standards in building market confidence by addressing blockchain issues. One concern is that these standardization attempts on the blockchain should align with relevant existing international industry standards. The integration of blockchain and existing industry standards and protocols, as well as the data storage over cloud systems, will be a key research issue.

From a long path on IloT, blockchain standardization will play a critical role in reshaping future technologies. The blockchain standardization should be able to provide guidance to developers and users on blockchain technologies.

IX. Conclusion

This paper provides a comprehensive and systematic review of the integration on blockchain and IloT platforms. We identify some key issues and challenges on this integration, from both blockchain and IloT platforms, separately. We then discuss the potential challenges with this kind of integration, as well as several key application areas in the industry. Finally, we provide several potential research directions on the integration of blockchain and IloT platforms.

REFERENCES


proof of burn, proof of capacity, and byzantines generals, bytcoin,” 2015.


