UNIVERSITY OF THE BASQUE COUNTRY

UPV/EHU

THESIS DISSERTATION

A Holistic DLT Architecture for Industry 4.0

Author:
Denis Ionut Stefanescu Ivan

Advisors:
Juan José Unzilla Galán
Leticia Montalvillo Mendizabal

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in the

Deparment of Communications Engineering

October 26, 2023

Typeset using LaTeX

ACKNOWLEDGEMENT

Al cerrar este importante capítulo de mi vida, quiero expresar mi profundo agradecimiento a aquellos que me han acompañado en este trascendental viaje.

En primer lugar, quiero agradecerles a mis supervisores, Juanjo Unzilla, Leticia Montalvillo y Patxi Galán-García, por su valiosa orientación, paciencia y constante apoyo. Sus conocimientos y sabiduría no solo han enriquecido este trabajo, sino que también han dejado una huella permanente en mi formación profesional y personal. Sin vosotros nada de esto habría sido posible, os estaré eternamente agradecido.

También quiero expresar mis agradecimientos a las instituciones que han hecho esto posible, la Universidad del País Vasco (UPV/EHU) y el centro de investigación IKERLAN, que me brindó las herramientas y recursos necesarios para llevar a cabo este proyecto. Gracias por creer en mis capacidades y brindarme oportunidades para crecer.

El apoyo de mis compañeros de IKERLAN ha sido una fuente inagotable de motivación. Quería darles las gracias a todos los compañeros del equipo de Ciberseguridad en Plataformas Digitales de IKERLAN (ZPD), por haber hecho de este proceso un recorrido enriquecedor y agradable. Sin embargo, hay algunos compañeros a quienes quisiera mencionar de manera especial por su inestimable apoyo durante este proceso: Aintzane Mosteiro, Xabier Saéz de Cámara, Gorka Abad, Adei Arias y Anhelina Kovach. Su pasión, visión y excelencia han sido el faro que ha guiado muchos de los momentos más desafiantes de mi camino. ¡Muchísimas gracias por todo!

Finalmente, a mi familia y amigos, quienes con su amor, comprensión y apoyo incondicional, han sido el sostén en los momentos más desafiantes. Este logro también es suyo.

Denis Ionut Stefanescu Ivan

Vitoria-Gasteiz, 2023

ABSTRACT

Industry 4.0, also recognized as the fourth industrial revolution, symbolizes an innovative phase of manufacturing. It is driven by emerging technologies like Artificial Intelligence (AI), the Internet of Things (IoT), Big Data, advanced robotics, augmented reality, cloud computing, and cybersecurity. The prime focus is to digitally transform and interconnect various production and logistics processes, thereby aiming to enhance productivity.

Within this Industry 4.0 paradigm, factories, production systems, and processes can self-monitor, self-adjust, and self-diagnose by utilizing real-time data analysis. In addition, products have the capability to interact with machinery, dictating the manufacturing process, while systems are able to self-learn ways to enhance production efficiency and quality.

Industry 4.0 harbors the potential to greatly optimize the manner in which goods and services are produced, offering improved efficiency, flexibility, and customization of products. However, this shift brings about several challenges, including concerns around data security and privacy, difficulties managing and scaling large volumes of data, standardization and compatibility problems due to the need to integrate varied data and systems, along with significant obstacles in executing automated processes. All these factors require substantial alterations in infrastructure and business practices.

To face the aforementioned challenges, an increasing number of field experts and researchers are exploring the potential utilization of Distributed Ledger Technologies (DLTs) within industrial settings. DLTs propose a groundbreaking approach to storing and sharing information, with decentralization being its distinguishing feature. Contrary to traditional databases or registries, which require a central entity for management and verification, DLTs empower secure and efficient recording, sharing, and verification of information among a network of users, thereby eliminating central authority. Broadly speaking, DLTs operate via nodes, each housing a copy of the ledger and assisting in transaction validation. DLTs stand out for their transparency, with all transactions being visible to every network participant, and their superior security, as every transaction requires network-wide consensus and once confirmed, cannot be altered or deleted. Consequently, compromising DLTs would require gaining control over the majority of

nodes, making it extremely challenging.

DLTs gained prominence due to blockchain technology, a specific variant of DLT. "Blockchain" refers to the unique data organization method of this technology. Rather than randomly grouping transactions, blockchain clusters them into blocks, linked chronologically, forming a blockchain.

Blockchain inherits various DLT attributes, such as decentralization, transparency, and security. It also possesses distinctive features. For instance, each block in the chain includes a cryptographic summary of its predecessor block, meaning that any alteration to a block changes its summary, subsequently affecting the next block and so on, thereby invalidating the entire chain. This characteristic renders the blockchain virtually immutable.

Blockchain's inaugural and most renowned application is Bitcoin, a cryptocurrency that reinvented the notion of digital currency by offering a secure, decentralized platform for financial transactions. However, since the creation of Bitcoin, the use of blockchain has expanded to a variety of applications, including smart contracts and supply chain tracking, highlighting its potential to reshape multiple economic sectors and societal facets.

Hence, in recent years, the intersection of blockchain technology and other DLTs with the evolving landscape of Industry 4.0 has garnered significant attention. A notable aspect of Industry 4.0 is its multi-layered, pyramid-like structure, often referred to as the industrial pyramid. At its foundation, data is generated by the Industrial IoT (IIoT) devices; as one ascends the pyramid, this data is then homogenized and processed within the industrial plant, and finally, at the top, it is harnessed for business decision-making and collaboration. While DLTs offer promising solutions to meet the demands of each of these layers, their implementation poses unique challenges. Furthermore, a comprehensive architecture that seamlessly integrates industrial layers with DLTs remains elusive.

Therefore, this thesis seeks to address the aforementioned Industry 4.0 challenges by proposing a DLT-based architecture that envelops the entirety of the industrial data lifecycle. Beginning at the machine level, where vast amounts of data are generated, the process transitions to the plant level for homogenization and processing, and ultimately

reaches the pinnacle where this data drives business logic and inter-company collaborations. However, existing DLTs face numerous performance and scalability issues, particularly at the foundational machine level, where prompt data processing is imperative. The middle, or plant level, requires sophisticated interoperability mechanisms, a domain yet to be fully matured. At the pyramid's top, the business level, there is a need for systems that can manage secure, automated digital contracts and maintain data confidentiality in an environment characterized by company interactions.

Therefore, this thesis focuses on designing a holistic DLT architecture for Industry 4.0 that effectively addresses the main challenges of the field and covers the whole cycle of the data. This architecture aims to create a secure and tamper-resistant environment for data, preserving its privacy and integrity, seeking to ensure efficient data standardization, promote seamless integration between diverse systems, and support automated processes through enhanced smart contracts with off-network data access capabilities. At the same time, this proposal is conceived to provide a scalable solution capable of handling the high volumes of real-time data generated within Industry 4.0, while also considering the energy efficiency and low monetary costs of the DLTs implementations. The scope of this architecture extends from data generation at the IoT level to its processing and use for business purposes at higher levels.

LABURPENA

Laugarren industria-iraultza, Industria 4.0 izenez ezagutzen dena, sortzen ari diren hainbat teknologia batzen dituen terminoa da, hala nola Gauzen Internet-a (ingelesez, Internet of Things (IoT)), Adimen Artifiziala (AA), Big Data, errealitate areagotua, robotika aurreratua, hodei-konputazioa eta zibersegurtasuna. Industria-iraultza honen helburu nagusia egungo industria-prozesuak hobetzea da, produkzio eta logistika prozesuen digitalizazioa eta interkonexioa sustatuz.

Industria 4.0-ren esparruan dauden fabriketan, sistemak eta ekoizpen-prozesuak haien kasa optimizatu, erregulatu eta diagnostikatu daitezke, datuak denbora errealean bilduz eta aztertuz. Produktuak makineriarekin ere komunikatu daitezke, nola fabrikatu behar diren beren kabuz zehazteko; sistemek, berriz, beren kabuz ikas dezakete ekoizpenaren eraginkortasuna eta kalitatea nola hobetu.

Industria 4.0-ren alderdi nabarmen bat geruza anitzeko egitura da, piramide baten antzekoa, askotan "piramide industriala" deitzen dena. Oinarrian, IoT Industrialeko (IIoT) gailuek datuak makina-mailan sortzen dituzte, eta piramidean gora egin ahala, datuak homogeneizatu eta prozesatu egiten dira lantegi barruan, azkenik, enpresa-mailako erabakiak hartzeko eta negozio-prozesuetan erabiltzeko.

Industria 4.0-k produktuak eta zerbitzuak ekoizteko modua eraldatzeko ahalmena dauka, produktuen eraginkortasun, malgutasun eta pertsonalizazio handiagoa ahalbidetuko lukeena. Hala ere, hainbat erronka ere planteatzen ditu, hala nola datuekin lotutako segurtasun eta pribatutasun arazoak, datu-bolumen handiak prozesatzeko beharraren ondorioz sortzen diren errendimendu eta hazkunde arazoak, datu eta sistema heterogeneoen interakzioek sortzen dituzten estandarizazio eta bateragarritasun arazoak, baita prozesu automatizatuak eta lotutako kostuak gauzatzeko ere. Guzti honek aldaketa esanguratsuak eskatzen ditu enpresen azpiegitura eta prozesuetan.

Aipatutako erronken ondorioz, gero eta ikertzaile eta aditu gehiagok aztertzen dute industria-inguruneetan Erregistro Banatuko Teknologiak (ingelesez Distributed Ledger Technologies (DLT)) aplikatzeko erak. DLTa informazioa gordetzeko eta partekatzeko modua irauli nahi duen teknologia da. DLTen faktore bereizgarri nagusia haien egitura deszentralizatua da. Transakzioak administratzeko eta egiaztatzeko erakunde zentralizatua behar duten erregistro edo datu-base tradizionalek ez bezala, DLTek hain-

bat parte-hartzaileren sare baten informazioa modu seguru eta eraginkorrean erregistratzea, partekatzea eta egiaztatzea ahalbidetzen dute, agintaritza zentral baten beharra ezabatuz. Oro har, DLTek nodoen bidez funtzionatzen dute, eta nodo horietako bakoitzak erregistroaren kopia bat gordetzen du, eta DLTan egiten diren transakzioak balioztatzen laguntzen du. DLTen ezaugarri nagusia gardentasuna da, sareko partehartzaile bakoitzak transakzio guztiak ikus ditzakelako. Gainera, DLTek oso segurtasun handia eskaintzen dute, transakzio bakoitza sareko partaideen arteko adostasunarekin berresten delako, eta transakzio bakoitza baieztatzen denean, ezin delako aldatu edo ezabatu. Ezaugarri horiek direla eta, DLTen segurtasuna bortxatzea oso zaila da, lan horrek sareko nodo gehienak kontrolatzea eskatuko lukeelako.

DLTak Blockchain teknologiari esker hedatu dira, Blockchain-a DLT-en mota konkretu bat da eta. "Blockchain" terminoak teknologia partikular honek datuak nola antolatzen dituen adierazten du. Transakzioak modu arbitrarioan multzokatu beharrean, blockchainetan datuak bloketan antolatu eta kronologikoki antolatzen dira, blokeen “kate” bat sortuz, hortik “blockchain” izena.

Blockchain-ek DLT-en ezaugarri asko hartzen dituzte, deszentralizazioa, gardentasuna eta segurtasuna, adibidez. Baina ezaugarri bereziak ere baditu, adibidez, kateko bloke bakoitzak aurreko blokearen laburpen kriptografiko bat dauka. Honi esker, bloke bateko edozein aldaketak haren laburpen kriptografikoa aldatuko luke, eta horrek, era berean, hurrengo blokea aldatuko luke, eta aldaketa guzti horiekin, kate osoa baliogabetuko litzateke. Ezaugarri horri esker, blockchain-a ia aldaezina da.

Blockchain-en lehen aplikazioa eta ospetsuena Bitcoin da, diru digitalaren kontzeptua irauli zuen kriptotxanpona, finantza-transakzioetarako plataforma deszentralizatua eta segurua eskaintzen baitzuen. Hala ere, Bitcoin asmatu zenetik, blockchain-entzako erabilera askoz gehiago aurkitu dira, kontratu adimendunetatik hasi eta hornidura-katearen jarraipeneraino, ekonomiaren eta gizartearen hainbat sektore eraldatzeko duen ahalmena erakutsiz.

Horregatik, azken urteotan, gero eta ikertzaile eta aditu gehiagok aztertu dute blockchain teknologia edo beste DLT batzuk Industria 4.0-ko inguruneetan ezartzeko aukera. Teknologia horien abantailak argiak dira etorkizuneko industrian dauden erronkei aurre egiteko orduan, baina teknologia horiek ezartzeak erronka ugari ditu.

Erronka garrantzitsuetako bat datuen prozesu osoa jarraitzen duen DLTetan oinarritutako arkitektura bat ezartzea da. Arkitektura honek datuak hainbat mailatan tratatu beharko ditu, makina-mailan datuaren sorreratik hasita, gero lantegi mailan prozesatu eta homogeneizatzeko, eta kanpo-mailako negozio-logiketarako ustiapenarekin amaituz, non enpresa oso ezberdinen arteko lankidetzak ugariak diren. Gainera, lehendik dauden DLTek errendimendu- eta eskalagarritasun-erronka ugari dituzte, eta horrek zaildu egiten du DLT hauek makina-mailan ezartzea; izan ere, makinetan datu ugari sortzen dira, eta arin prozesatu behar dira. Lantegiari dagokionez, DLTek haien artean lan egiteko mekanismo aurreratuak behar dituzte, eta alderdi hori gaur egun ez dago behar bezain aurreraturik. Datuak negozio-mailan ustiatzerakoan, enpresa desberdinen arteko elkarrekintzak dauden inguruneetan, kontratu digital seguruak eta automatizatuak egitea ahalbidetuko duten mekanismoak ezarri behar dira, baita datuen pribatutasuna bermatuko duten beste mekanismo batzuk ere. Azkenik, garrantzitsua da DLT teknologiei lotutako kostuak ere kontuan izatea; izan ere, DLTetan egiten diren transakzioen balidazioen kostuak altuak izan daitezke, baita balidazio prozesu hauek erabiltzen duten energiaren kostea ere.

Beraz, tesi honen ardatz nagusia Industria 4.0rako DLT arkitektura integrala diseinatzea da, eremu horretan dauden erronka nagusiei eraginkortasunez erantzungo diena, eta prozesu osoa barne hartuko duena, datua makina-mailan sortzen denetik negozio-mailan prozesatu eta ustiatzen den arte. Arkitektura honek segurua den eta datuen manipulazioa ekidituko duen ingurunea sortzea du helburu, datuen pribatutasuna eta osotasuna babestuz, datuen estandarizazio eraginkorra bermatuz, sistema anitzen arteko integrazio egokia sustatuz, eta saretik kanpoko datuak eskuratzeko ahalmena duten kontratu adimendun hobetuen bidez prozesu automatizatuak lagunduz. Proposamen hau Industria 4.0ren barruan sortutako denbora errealeko datu-bolumen altuak kudeatzeko gai izango den irtenbide eskalagarri bat lortzeko sortu da, ingurumenaren gaineko eragina, eraginkortasun energetikoa eta DLTei lotutako kostuak ere kontuan hartzen dituena. Bide berean, arkitektura honen irismena makinetan datuak sortzen direnetik, prozesatu eta helburu komertzialetarako erabiltzen diren arte hedatzen da.

Lehen aipatutako arazoei erantzuteko eta tesiaren helburuak betetzeko, hainbat ekarpen egin dira.

Lehenik eta behin, Industria 4.0ren eremua aztertu da, tesiaren abiapuntu modura. Egungo egoera hau automatizazio-piramidearen eredua jarraituz egituratu da, manufakturaren sektorean erabili ohi dena, eta lau etapatan banatzen da, industria-eragiketen maila desberdinak sinbolizatzen dituztenak. Definitutako egoera hau automatizazio-piramidearen bilakaera gisa definitzen da, industria 4.0rekin eta industria hori osatzen duten teknologia disruptiboekin loturiko erronka berriei erantzungo diena; kasu honetan, DLTak bereziki.

Lehenik, makinaren maila dago, automatizazio-piramidearen eremu eta kontrol etapetatik ateratakoa. Maila honetan, industriako gailuak optimizatu nahi dira, hala nola makinak eta sentsoreak, haien errendimendua eta eraginkortasuna hobetzeko, baita kontrol-sistemekin integratzeko ere.

Ondoren, ekoizpen-lerroaren maila aurkezten da, automatizazio-piramidearen gainbegiratze-mailan oinarrituta. Maila honetan, ekoizpen-lerro guztiak hobetu nahi dira, hainbat elementu koordinatzen denbora hilak murrizteko, eta produktibitatea eta produktuaren kalitate handiena bermatzeko.

Hirugarren maila lantegiarena da, automatizazio-piramidearen "Management Execution System" (MES) mailarekin lerrokatua. Hemen, helburua da industria-lantegi osoen eraginkortasuna hobetzea da, ekoizpen-lerroen datuak, baliabideen erabilera eta hondakinen murrizketa administratuz, beste hainbat lanen artean.

Azkenik, partzuergo mailak, piramidearen Enterprise Resource Planning (ERP) mailan oinarrituta, hainbat industria-lantegi Industria 4.0ren negozio partzuergo bakar baten batzeko aukera aztertzen du. Maila honek lankidetza handiagoa, datuen trukea eta baliabideen optimizazioa sustatzen ditu.

Hurrengo urrats gisa, tesi honetan literaturaren berrikuspen sistematiko bat egiten da (ingelesez Systematic Literature Review (SLR)), IoT-arekin lotutako aplikazioetan DLT arkitekturak erabiltzeari buruzkoa, Industria 4.0n batez ere. DLTek baliabide mugatuak dituzten gailuekin lan egiteko egoeretan jasaten dituzten errendimendu eta eskalabilitate murrizketak ikusita, ikerketaren zati handi bat egungo lanetan zentratu da, baldintza horietan eraginkorrak izango diren konponbideak garatzera bideraturik. SLR erabiliz, sakon aztertu dira arkitektura horiek, haien ezaugarriak eta ebaluazioak nabarmenduz.

Zehazki, ikerketa honek segurtasunaren, pribatutasunaren, eraginkortasunaren eta

eskalagarritasunaren erronkekin egiten du lan batez ere. DLTak erabiltzearen hainbat onura nabarmentzen dituen arren, deszentralizazioa, iraunkortasuna edota IoT-ren segurtasuna eta eskalagarritasuna hobetzeko auditoretza erabiltzearen aukera, ikerketak baliabide mugatuak dituzten inguruneetan DLTak ezartzeko zailtasunak ere azaltzen ditu. IoT gailuak ugari diren inguruneekin bateragarriak diren arkitektura arinen garrantzia azpimarratzen direlarik.

Ikerketa honek tesi-proiektuan DLTetan oinarritutako arkitektura integral bat garatzeko oinarri bat sortu du, dauden arkitekturen mugak gainditu eta oraindik aztertu gabeko ikerketa-aukerak aprobetxatzea ahalbidetuz, datuen bizi-ziklo osoa kontuan harturik, eta ez soilik IoT makinei eta gailuei dagokizkien atalak.

Azkenik, tesi honen funtsezko ekarpen gisa, DLTetan oinarritutako geruza ugariko arkitektura aurkezten da, Industria 4.0ko benetako kasu baten datuen kudeaketa eta segurtasuna hobetzeko diseinatu dena. Arkitektura hiru geruzatan egituratzen da, industria-prozesuan zehar sistematikoki antolatuak, makineriaren lanetatik hasi eta goi-mailako enpresa-erabakietaraino.

Lehen geruza, "Data Source Layer", ekoizpen-lerroaren mailan kokatzen da, eta gailu industrialek sortzen dituzten datuez arduratzen da. Geruza honen funtzio nagusia datuak denbora errealean jaso, biltegiratu eta kudeatzea da, datuen osotasuna ziurtatuz eta makinen mailako jardueren erregistro bortxaezina eskainiz.

Bigarren geruza, "Bridge Layer", lantegi mailan kokatzen da. Geruza honetan, lehen geruzako datuen agregazioa egiten da, eta lantegi baten barruan dauden produkzio-lerro desberdinen arteko komunikazioa eta datu-trukea errazten du. Horrela, fabrikazio-ekosistema integratuagoa eta eraginkorragoa sortzen laguntzen du.

Hirugarren geruzak, "Business Layer", partzuergo eta negozio mailan jarduten du. Geruza hau eragile bakoitzaren industria-lantegietako datuak prozesatzeko, aztertzeko eta kudeatzeko arduraduna da. Enpresen erabaki estrategikoak hartzen eta beharrezko ikuspuntuak lortzen laguntzen du. Gainera, datuen segurtasuna, pribatutasuna eta trazabilitatea bermatzen ditu, partzuergoko kideek eta kanpoko interesdunek akordio automatikoak ahalik eta konfiantza handienarekin egin ditzaten.

Azkenik, arkitektura ingurune desberdinetan balioztatu da, bai simulatuetan, baita ingurune errealistago batean, IKERLAN zentro teknologikoaren eta Fagor Automa-

tion industria-enpresaren lankidetzarekin garatzen den erabilera errealeko kasu bati erantzuna emanez.

Kapitulu honetan proposatutako arkitekturak datuak kudeatzeko sistema integral, seguru eta eraginkor bat garatzen du, industri ekosistema osoa kontutan hortzen duena. Integrazio horrek komunikazioa eta informazio-trukea errazten du. Ondorioz, fabrikazio-ingurune interkonektatuagoa eta seguruagoa sortzen da. Era honetan, Industria 4.0-ren funtsezko erronkei erantzuten zaie eta enpresei eraldaketa prozesurako beharrezko pausuak errazten zaizkie.

RESUMEN EJECUTIVO

La cuarta revolución industrial, conocida como Industria 4.0, es un término que combina una serie de tecnologías emergentes como el Internet de las Cosas (en inglés, Internet of Things (IoT)), la Inteligencia Artificial (IA), el Big Data, la realidad aumentada, la robótica avanzada, la computación en la nube o la ciberseguridad para mejorar los procesos industriales de la actualidad. Por tanto, esta revolución industrial está enfocada en la digitalización y la interconexión de los procesos de producción y logística, con el objetivo prinicipal de lograr una mayor productividad.

Las fábricas que se encuentran dentro del marco de Industria 4.0, los sistemas y los procesos de producción pueden auto-optimizarse, autorregularse y autodiagnosticarse a través de la recopilación y el análisis de datos en tiempo real. Los productos también pueden comunicarse con la maquinaria para establecer por sí mismos cómo deben ser fabricados, mientras que los sistemas pueden aprender por sí mismos a mejorar la eficiencia y la calidad de la producción.

Un aspecto destacado de la Industria 4.0 es su estructura multicapa, semejante a una pirámide, a menudo referida como la pirámide industrial. En su base, los datos son generados a nivel de máquina por los dispositivos del IoT Industrial (IIoT); ascendiendo en la pirámide, los datos son homogeneizados y procesados dentro de la planta industrial y, finalmente, son utilizados para la toma de decisiones empresariales y los procesos de negocio.

La Industria 4.0 tiene el potencial de transformar la forma en que producimos bienes y servicios, lo que permitiría una mayor eficiencia, flexibilidad y personalización de los productos. Sin embargo, también plantea una serie de desafíos, como problemas de seguridad y privacidad relacionados con los datos, problemas de rendimiento y escalabilidad debido a la necesidad de procesamiento de grandes volúmenes de datos, problemas de estandarización y compatibilidad debido a la necesidad de interacción de datos y sistemas heterogéneos, así como retos importantes en la ejecución de procesos automatizados y los costes asociados. Todo esto conlleva la necesidad de cambios significativos en la infraestructura y los procesos de las empresas.

Debido a los retos mencionados anteriormente, cada vez más investigadores y expertos en el ámbito han estudiado la posible aplicación de tecnologías de registro dis-

tribuido (en inglés Distributed Ledger Technologies (DLT)) en entornos industriales. Las DLTs son una tecnología que pretende revolucionar la manera en que se almacena y se comparte la información. El principal factor distintivo de las DLTs es su estructura descentralizada. A diferencia de los registros o bases de datos tradicionales, que requieren una entidad centralizada para administrar y verificar las transacciones, las DLTs permiten que la información se registre, comparta y verifique de forma segura y eficiente entre una red de participantes, eliminando la necesidad de una autoridad central. En términos generales, las DLTs funcionan a través de nodos, cada uno de los cuales almacena una copia del registro y colabora en la validación de las transacciones. Las DLT se caracterizan por su transparencia, ya que cada participante en la red puede ver todas las transacciones. Además, las DLTs ofrecen una seguridad muy alta porque cada transacción es confirmada por consenso entre los participantes en la red, y una vez que se confirma una transacción, no puede ser modificada ni eliminada. Estos aspectos hacen que las DLTs sean extremadamente difíciles de comprometer, ya que esta labor requeriría el control de la mayoría de los nodos en la red.

Las DLTs se han popularizado gracias a la tecnología blockchain, que es una forma específica de DLT. El término "blockchain" se refiere a cómo esta tecnología particular organiza los datos. En lugar de agrupar las transacciones de manera arbitraria, las transacciones en una blockchain se agrupan en bloques, que se encadenan cronológicamente. Esto crea una cadena de bloques: "blockchain".

Las blockchain heredan muchas de las características de la DLT, como la descentralización, la transparencia y la seguridad. Pero también tiene características únicas. Por ejemplo, cada bloque en la cadena contiene un resumen criptográfico del bloque anterior. Esto significa que cualquier cambio en un bloque alteraría su resumen, lo que a su vez alteraría el bloque siguiente, y así sucesivamente, invalidando toda la cadena. Esta característica hace que el blockchain sea prácticamente inalterable.

La primera y más famosa aplicación del blockchain es el Bitcoin, una criptomoneda que revolucionó el concepto de dinero digital al proporcionar una plataforma descentralizada y segura para las transacciones financieras. Sin embargo, desde la invención de Bitcoin, el blockchain ha encontrado muchas más aplicaciones, desde contratos inteligentes hasta el seguimiento de la cadena de suministro, demostrando su potencial

para transformar diversos sectores de la economía y la sociedad.

Es por esto que, en los últimos años, cada vez más investigadores y expertos han estudiado la posibilidad de implementar la tecnología blockchain u otras DLTs en entornos de Industria 4.0. Aunque las ventajas de estas tecnologías son claras a la hora de resolver los retos presentes en la industria del futuro, su implantación no está exenta de numerosos retos propios.

Un reto importante que surge es el hecho de lograr implementar una arquitectura basada en DLTs que cubra todo el proceso del dato, empezando por el momento en el cual este se genera a nivel de máquina, siguiendo por su posterior procesamiento y homogeneización a nivel de planta, y finalizando con su explotación para lógicas de negocio a nivel externo, donde abundan las colaboraciones entre empresas muy diferentes. Además, las DLT existentes poseen numerosos retos de rendimiento y escalabilidad que dificultan su implementación a nivel de máquina, donde se generan grandes cantidades de datos que necesitan ser procesados de forma ágil. A nivel de planta, nos encontramos con el hecho de que las DLT necesitan mecanismos avanzados de interoperabilidad, siendo este un aspecto que en la actualidad no está lo suficientemente avanzado. A la hora de explotar los datos a nivel de negocio, se deben implementar mecanismos que permitan la realización de contratos digitales seguros y automatizados, así como otros mecanismos que garanticen la privacidad del dato en entornos donde existen interacciones entre numerosas empresas distintas. Finalmente, también es importante tomar en cuenta los costes asociados a las tecnologías DLT, ya que estas pueden acarrear costes importantes a la hora de validar transacciones, o costes en términos de energía a la hora de realizar dicha validación.

Por tanto, esta tesis se centra en diseñar una arquitectura DLT integral para la Industria 4.0 que aborde de manera efectiva los principales desafíos presentes en dicho ámbito, y que cubra todo el proceso, desde que el dato se genera a nivel de máquina, hasta que se procesa y explota a nivel de negocio. Esta arquitectura pretende crear un entorno seguro y resistente a manipulaciones para los datos, preservando su privacidad e integridad, buscando garantizar una eficiente estandarización de datos, promover una integración idónea entre sistemas diversos, y respaldar procesos automatizados a través de contratos inteligentes mejorados con capacidades de acceso a datos fuera de la red.

Al mismo tiempo, esta propuesta se concibe para proporcionar una solución escalable capaz de gestionar los altos volúmenes de datos en tiempo real generados dentro de la Industria 4.0, considerando además el impacto medioambiental, la eficiencia energética y los costes asociados de las DLTs. Por tanto, el alcance de esta arquitectura se extiende desde la generación de datos en las máquinas hasta su procesamiento y utilización para fines comerciales.

Para dar respuesta a los problemas mencionados anteriormente y cumplir los objetivos de la tesis, se han realizado varias contribuciones.

Primero se examina el ámbito de la Industria 4.0 que sirva como escenario de partida de la tesis. Este escenario se estructura siguiendo el modelo de la pirámide de automatización, comúnmente utilizado en el sector de la manufactura, y se divide en cuatro etapas que simbolizan distintos grados de operaciones industriales. El escenario definido se define como una evolución de la pirámide de automatización que dé respuesta a los nuevos retos relacionados con la Industria 4.0 y las tecnologías disruptivas que forman parte de ella, siendo en este caso principalmente las DLTs.

Primero se encuentra el nivel de máquina, extraído de las etapas de campo y control de la pirámide de automatización. En este nivel, se busca la optimización de componentes industriales, como máquinas y sensores, con el fin de mejorar su rendimiento y eficacia, así como su integración con sistemas de control.

Después se presenta el nivel de línea de producción, inspirado en el nivel de supervisión de la pirámide de automatización. En este nivel, se persigue el mejoramiento de las líneas de producción en su totalidad, coordinando múltiples elementos para reducir tiempos muertos y garantizar la máxima productividad y calidad del producto.

El tercer nivel es el de planta, alineado con el nivel de Management Execution System (MES) de la pirámide de automatización. Aquí, el objetivo es mejorar la eficiencia de las plantas industriales completas, administrando desde los datos de las líneas de producción hasta la utilización de recursos y la disminución de desechos.

Finalmente, el nivel de consorcio, basado en el nivel Enterprise Resource Planning (ERP) de la pirámide, explora la posibilidad de unir varias plantas industriales en un consorcorcio de negocios de la Industria 4.0. Este nivel promueve una mayor colaboración, el intercambio de datos y la optimización de recursos.

Como siguiente paso, en esta tesis se realiza una revisión sistemática de la literatura (en inglés Systematic Literature Review (SLR)) sobre la utilización de arquitecturas DLT en aplicaciones relacionadas con el IoT, con especial enfoque en la Industria 4.0. En vista de las restricciones de rendimiento y escalabilidad de las DLT en situaciones que involucran dispositivos con recursos limitados, gran parte de la investigación se ha centrado en los trabajos actuales que se han enfocado en desarrollar soluciones que sean eficaces en estas condiciones. La SLR analiza en profundidad dichas arquitecturas, resaltando sus características y evaluaciones.

Específicamente, el estudio se enfoca en retos como la seguridad, la privacidad, la eficiencia y la escalabilidad. Mientras que destaca los beneficios de usar DLTs, como la descentralización, la persistencia y la posibilidad de auditoría para mejorar la seguridad y la escalabilidad de IoT, el estudio también reconoce las dificultades para implementar DLTs en entornos con recursos limitados, subrayando la importancia de las arquitecturas ligeras que sean compatibles con los entornos donde abundan los dispositivos IoT.

Este estudio otorga a esta tesis una base para el desarrollo de una arquitectura integral basada en DLTs que pueda superar las limitaciones de las arquitecturas existentes y aprovechar las oportunidades de investigación no exploradas aún, teniendo en cuenta todo el ciclo de vida de los datos, no solo la porción relativa a las máquinas y dispositivos IoT.

Finalmente, como aportaciones clave de esta tesis, se presenta una arquitectura de múltiples capas basada en DLTs, que ha sido diseñada para mejorar la gestión de datos y la seguridad en un escenario de Industria 4.0. La arquitectura se estructura en tres capas distintas, dispuestas sistemáticamente a lo largo del proceso industrial, desde la operación de la maquinaria hasta las decisiones empresariales de alto nivel.

La primera capa, "Data Source Layer", se ubica a nivel de línea de producción. Se encarga de los datos generados por los dispositivos industriales. La función principal de esta capa es capturar, almacenar y gestionar los datos en tiempo real, asegurando la integridad de los datos y proporcionando un registro inviolable de las actividades a nivel de máquina.

La segunda capa, "Bridge Layer", se sitúa a nivel de planta. En esta capa se realiza el agregado de los datos de la primera capa y facilita la comunicación y el intercambio

de datos entre diferentes líneas de producción dentro de una planta. De este modo, contribuye a la creación de un ecosistema de fabricación más integrado y eficiente.

La tercera capa, "Business Layer", opera a nivel de consorcio o negocio. Esta capa es responsable del procesamiento, análisis y gestión de datos procedentes las plantas industriales de cada actor. Apoya las decisiones empresariales estratégicas y ayuda a obtener las percepciones pertinentes. Además, garantiza la seguridad, la privacidad y la trazabilidad de los datos, permitiendo que los miembros del consorcio y los interesados externos ejecuten acuerdos automáticos con la máxima confianza posible.

Finalmente, la arquitectura se valida en diferentes entornos, tanto simulados como en un entorno más realista, dando respuesta a un caso de uso real que se desarrolla con la colaboración del centro tecnológico IKERLAN junto a la empresa industrial Fagor Automation.

Por tanto, la arquitectura propuesta en este capítulo desarrolla un sistema integral, seguro y eficiente para el manejo de datos que cubre todo el ecosistema industrial. Esta integración facilita una comunicación e intercambio de información fluidos, dando lugar a un entorno de fabricación más interconectado y seguro, que da respuesta a los retos fundamentales de la Industria 4.0 y facilita la adpoción de ésta en las empresas.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TERMS AND ABBREVIATIONS

AES	Advanced Encryption Standard
AI	Artificial Intelligence
AML	Anti Money Laundering
API	Application Programming Interface
AWS	Amazon Web Services
BCoT	Blockchain of Things
BFT	Byzantine Fault Tolerance
BIoT	Blockchain for IoT
CH	Cluster Head
CNC	Computer Numerical Control
DAG	Directed Acyclic Graph
DAO	Decentralized Autonomous Organization
DCS	Distributed Control System
DDoS	Distributed Denial of Service
DLT	Distributed Ledger Technology
DLTs	Distributed Ledgers
DoS	Denial of Service
DSR	Design Science Research
ECDSA	Elliptic Curve Digital Signature Algorithm
EdDSA	Edwards-curve Digital Signature Algorithm
ERP	Enterprise Resource Planning
FL	Federated Learning
FPGA	Field Programmable Gate Array
HMI	Human-Machine Interface
IIoT	Industrial IoT
IoD	Internet of Drones
IoT	Internet of Things

IPFS	Interplanetary File System
IT	Information Technology
JSON	JavaScript Object Notation
KYC	Know Your Customer
M2M	Machine-to-Machine
MCMC	Markov Chain Montecarlo
MES	Manufacturing Execution System
NIST	National Institute of Standards and Technology
OEE	Overall Equipment Effectiveness
P2P	Peer-to-Peer
PBFT	Practical Byzantine Fault Tolerance
PLC	Programmable Logic Controller
PoA	Proof of Authority
PoAc	Proof of Activity
PoB	Proof of Burn
PoC	Proof of Capacity
PoET	Proof of Elapsed Time
PoR	Proof of Reputation
PoS	Proof of Stake
PoW	Proof of Work
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RAM	Random Access Memory
RFID	Radio Frequency Identification
RTU	Remote Terminal Unit
SCADA	Supervisory Control And Data Acquisition
SDK	Software Development Kit
SLR	Systematic Literature Review
SSD	Solid State Drive
TLS	Transport Layer Security
VP	Validation Period

WSN Wireless Sensor Network

Chapter 1

Introduction

This chapter provides the reader with a comprehensive overview of the thesis. First, Section 1.1 contextualizes the thesis, while Section 1.2 introduces the motivation and research questions to be addressed. Finally, Section 1.3 describes the followed research methodology and Section 1.4 presents the outline of the thesis.

1.1 Context and General Problem Overview

Industry 4.0 - Introduction and Problem Statement

The Fourth Industrial Revolution, known as Industry 4.0, represents the progression of automation within manufacturing and various other sectors. This concept is marked by the implementation of technologies such as the Internet of Things (IoT), Artificial Intelligence (AI), cloud computing, and data analytics to boost productivity, efficiency, and innovation [1]. A primary factor propelling Industry 4.0 is the escalating digitalization of manufacturing and other industries. This process entails employing sensors, data analytics, and additional technologies to collect and examine real-time data from machinery, procedures, and products. The analyzed data is then utilized to refine operations, enhance product quality, and reduce associated costs [2].

Having grasped the magnitude and scope of Industry 4.0, it becomes imperative to position it in the context of how businesses and industries structure their operations in terms of automation. While Industry 4.0 includes an integrated and interconnected vision of operations, it is the automation pyramid model [3] that offers a structured framework to understand how these technological advancements are implemented and organized in practice.

The automation pyramid (depicted in Figure 1.1) provides a hierarchical depiction of the various levels of control and management in an industrial setting:

- Field Level: The foundational layer, encompassing physical components such as sensors, actuators, and other hardware that directly interacts with the production environment.
- Control Level: This tier consists of devices like Programmable Logic Controller (PLC) and Remote Terminal Unit (RTU). They gather data from the field level and make real-time decisions based on pre-established conditions.
- Supervisory Level: Dedicated to monitoring and controlling a myriad of interconnected devices, this level affords operators an overarching view of the production environment and allows for remote control of various processes.
- Planning Level: This layer addresses planning and management tasks, encompassing functions like production scheduling, inventory management, and quality assurance.
- Management Level: Sitting atop the pyramid, this level deals with overarching business functions, such as Enterprise Resource Planning (ERP) systems, which integrate and manage core business processes.

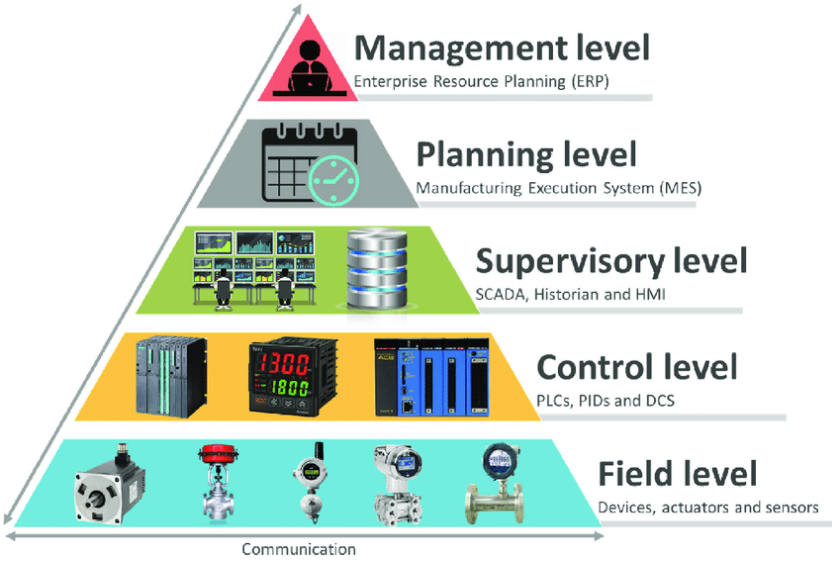

Fig. 1.1 Automation pyramid diagram [4]

However, despite its revolutionary promise, the materialization of Industry 4.0 remains largely unrealized. At its core, the issue is not just the complexity of integrating new technologies into existing infrastructure; it is an intricate web of challenges that

hinder its full-scale adoption. Companies find themselves navigating a maze of financial, technical, and operational barriers that make the leap from traditional industrial practices to a fully-realized Industry 4.0 model seem like a distant dream rather than an imminent reality. As a result, for many organizations, Industry 4.0 remains more of a theoretical ideal rather than a practical initiative.

While companies may flirt with aspects of Industry 4.0 —perhaps by integrating some automated systems or employing big data analytics— the full vision, replete with its transformational power, remains elusive. Consequently, the gap between what Industry 4.0 could be and what it currently is continues to widen, raising questions about whether this industrial revolution may, in fact, be more of an aspiration than an achievable goal for many.

Therefore, the **problem statement of this thesis** can be defined as follows: *"The revolutionary promise of Industry 4.0 remains largely unrealized due to significant barriers to its adoption"*.

The next subsection will discuss the main challenges that cause the problem outlined above.

Industry 4.0 Challenges

While Industry 4.0 offers numerous benefits and opportunities for businesses, it also brings about a variety of challenges in several dimensions: performance and scalability, data standardization and interoperability, automation and costs. However, the most concerning challenges are related to the cybersecurity [5] dimension. Specifically, the increasing inter-connectivity of devices and systems, as well as the widespread adoption of digital technologies, introduces new vulnerabilities and attack surfaces for malicious actors to exploit [6] [7] [8].

The present dimensions and their specific subsequent challenges in detail are as follows:

Cybersecurity: The practice of protecting digital infrastructure, networks, and data from unauthorized access, damage, and threats.

Cybersecurity challenges include several aspects as follows:

- Achieving decentralization of trust: Centralization creates significant vulnerabilities. If the central authority is compromised, the entire system can be brought down. This aspect demands a solution that provides a robust level of security and trust without centralization, enabling a resilient system with no single point

of failure. In a world that has conventionally been centralized, the burden of trust has typically rested upon certain entities or intermediaries [9], known for their reputation and stringent security measures. They form a network of trusted nodes, validating transactions and safeguarding the integrity of the system as a whole. This method, although effective to a certain degree, places a heavy reliance on these few entities.

- Building immutable records: There is a need for a system where records of transactions or events are securely stored in a way that is unchangeable and verifiable. The absence of this feature opens the door for fraud and manipulation. The task of keeping records safe and unalterable has traditionally been a meticulous process of logging and auditing. Organizations have used database management systems that capture all changes to records in a comprehensive, traceable manner. Additionally, they have utilized cryptographic techniques as an additional layer of security, helping to flag any unauthorized modifications. Yet, these processes can be labor-intensive and still may not guarantee total immunity from tampering [10].

- Access control: Despite numerous advances in cybersecurity, controlling who has access to specific data and who can perform actions within a system remains a significant challenge. Balancing accessibility and privacy while maintaining security is a complex problem that needs to be addressed. To regulate access, stringent policies and protocols are put into place, bolstered by robust identity and access management solutions [11]. These solutions offer tools for defining who gets to access what data and perform particular actions within a system. Employing two-factor authentication, role-based access control, and comprehensive activity logging are among the strategies used to enhance security, providing a fine-grained control over access privileges.

Performance and scalability: A system's ability to maintain efficient functionality and adapt its capacity as user demand or network activity increases, handling high volumes of data and transactions without performance degradation.

Performance and scalability challenges include several aspects as follows:

- Large data processing: As data volumes increase exponentially, traditional centralized systems struggle to keep up. Efficiently distributing data processing tasks across a network can reduce bottlenecks and improve system performance. Distributed data processing, for instance, has been facilitated by distributed computing frameworks [12]. These frameworks break down complex computational

tasks into smaller, manageable units, distributing them across multiple computers or servers for concurrent processing. This enables companies to handle vast volumes of data and compute-intensive tasks, alleviating the pressure on any single machine.

- Infrastructure growth: As user demand and network activity increase, systems must adapt to handle the load effectively. The ability to scale up and accommodate this growth without performance degradation is a key challenge. To address the inevitable need for growth and expansion, scalable architecture designs have been adopted. Elastic computing solutions allow systems to automatically scale up or down, based on the demand, facilitating an efficient use of resources. However, ensuring seamless scalability without performance degradation remains a complex task [13].

Data standardization and interoperability: The establishment of consistent data formats and communication protocols, facilitating seamless interaction between diverse systems, cross-system transactions, and data portability, while promoting efficient multi-party collaboration.

Data standardization and interoperability challenges include several aspects as follows [14]:

- Inconsistent data formats: With the variety of data types and structures used by different systems, establishing a consistent, standardized data format that enables seamless interaction is vital. Lack of such standards can hinder interoperability and complicate data analysis.

- Heterogeneous systems integration: Many systems today are built independently with unique data structures and formats. These differences can pose significant challenges when integration is required, potentially leading to data loss or misinterpretation.

- Data portability: The ability to move data securely and accurately from one system to another is a key requirement in today's interconnected world. Any loss, distortion, or misinterpretation of data during this process can lead to serious consequences.

- Multi-party collaboration: Establishing a system that allows multiple entities to collaborate securely and efficiently is challenging. This is especially true in environments where each entity may have its own unique processes and systems.

Automation: The use of technology to perform tasks with minimal human intervention, focusing on aspects such as automated contract enforcement, ensuring transparency in automated processes, and maintaining regulatory compliance within these automated systems.

Automation challenges include several aspects as follows:

- Automated contract enforcement: Enforcing contractual agreements in a digital space is a challenging task, especially when it needs to be carried out without human intervention. Implementing such a feature requires a solution that is not only technically feasible but also legally sound. Digital contracts, for example, need to be enforced automatically without requiring manual intervention. To address this, traditional systems have employed business rule engines that automate decision-making based on predefined rules and policies [15]. Despite these efforts, ensuring transparency in automated systems and maintaining regulatory compliance continue to be daunting tasks.

- Transparency in automated systems: Trust in automated systems is vital. However, ensuring transparency in automated decision-making processes, so all stakeholders understand how decisions are made, remains a significant challenge.

- Regulatory compliance: Automation can lead to significant efficiencies, but ensuring these systems comply with relevant regulations can be complex and time-consuming. Non-compliance can result in hefty fines and reputational damage.

Costs: The financial investment required to implement and maintain Industry 4.0 technologies, focusing on the costs of software licenses, hardware acquisition, employee training, and ongoing operational expenses.

Cost challenges include several aspects as follows:

- Software licenses and hardware acquisition: Implementing Industry 4.0 solutions often requires significant investment in software licenses and hardware equipment. This investment could be particularly high for specialized software solutions and state-of-the-art hardware. The high costs often present a barrier for small and medium-sized enterprises [16].

- Ongoing operational expenses: Beyond initial implementation, the costs of maintaining and upgrading Industry 4.0 technologies can be high. This includes subscription fees for software solutions, maintenance contracts for hardware, and costs related to data storage and management.

Thus, in the context of Industry 4.0, according to the latest research, **the key dimensions and their subsequent challenges that cause the problems that this thesis is trying to tackle** are as follows: *"Cybersecurity, performance and scalability, data standardization and interoperability, automation and costs".*

Building upon the key challenges identified, the following subsection will offer a comprehensive analysis of their consequences within the Industry 4.0 landscape and exhibit the significance of resolving the problem.

Consequences of the Industry 4.0 Challenges

Failure to fully adopt Industry 4.0 technologies has a significant impact that extends far beyond just the companies involved. One of the most immediate impacts is the loss of productivity [17]. In an era where efficiency and automation are crucial, manufacturing companies that do not integrate these advanced technologies risk falling behind in terms of operational efficiencies. This puts them at a competitive disadvantage, which is another significant repercussion.

The issue also extends to human capital, since this leads to a drain of highly qualified personnel [18]. The departure of such talent further contributes to the decrease in productivity and competitiveness, creating a vicious cycle that is difficult to break.

Finally, customer dissatisfaction becomes a major concern [19]. In a world where customers are becoming increasingly accustomed to the benefits of technology, such as rapid delivery times, high customization, and superior quality, failure to adopt Industry 4.0 technologies can be acutely felt by the end consumer. This dissatisfaction not only impacts revenue but also damages the brand image of the companies, making it more challenging to attract and retain customers in the long term.

In sum, **the consequences of failing to address the barriers to Industry 4.0 adoption** are as follows: *"Loss of productivity and competitiveness, skilled labor, and customer satisfaction".*

Given the severe consequences associated with the failure to tackle the challenges in Industry 4.0, it becomes imperative to explore innovative solutions that can fill these gaps. Emerging technologies offer the promise of overcoming such obstacles, setting the stage for the next era of industrial advancement.

DLTs for Industry 4.0

As previously stated, the emergence of Industry 4.0 introduces a plethora of challenges requiring immediate and innovative solutions. Traditional technologies have played a pivotal role in mitigating some of these pressing concerns.

Centralized databases and cloud solutions have served as the backbone for data storage and management. Despite their utility, these systems pose risks related to single points of failure [20]. They can also suffer from latency issues, limiting their effectiveness in real-time data processing. Data analytics have facilitated meaningful insights from vast data pools but often require significant computational resources and may lack the real-time processing capabilities needed for instantaneous decision-making. Standard encryption techniques have provided a layer of security, but they are often not robust enough to fend off increasingly sophisticated cyber-attacks [21].

It is clear that while considerable strides have been made in addressing the present Industry 4.0 challenges, there are still areas where existing solutions fall short, paving the way for newer, potentially more efficient technologies, such as Distributed Ledger Technology (DLT). Thus, in the face of these challenges, Distributed Ledgers (DLTs) are being seen as potentially disruptive solutions that can address these hurdles and unlock the full potential of the Industry 4.0 revolution [22] [23] [24].

DLTs allow for secure and transparent recording of transactions and data. They encompass a network of nodes that cooperate to validate and record transactions on a shared, immutable decentralized ledger. In most DLTs, such as blockchain, which is the most relevant DLT at the time, each entry on the ledger is cryptographically linked to its predecessor, creating a secure and tamper-resistant record of all transactions. Any alteration to a single entry would necessitate changing all subsequent ones, a task that is highly unfeasible [25].

One of the core advantages of DLTs is their inherent decentralization, providing resilience against tampering and fraudulent activities [26]. Beyond that, DLTs can also execute automated agreements or processes, known as "smart contracts". Smart contracts are instrumental in tackling the Industry 4.0 challenges, particularly in cybersecurity, interoperability and automation. Smart contracts automate transaction execution and monitoring, helping to prevent time loss and errors induced by manual intervention, thus enhancing transparency, efficiency and security [27].

On the other hand, interoperable DLTs can enable seamless compatibility and data standardization and provide security throughout the whole data lifecycle [28]. This is due to the fact that this shared, decentralized database can serve as a "common language" across different systems and platforms, drastically simplifying the process of data standardization and ensuring consistency of information, which can dramatically

expand the scope and capabilities of these systems by enabling the creation of complex, interchain applications and fostering a more integrated and collaborative ecosystem of distributed networks.

While the potential benefits of applying DLTs to address challenges in Industry 4.0 are apparent, there are several obstacles that need to be overcome for successful implementation [29]. Industry 4.0 environments are complex, and data security is crucial at every stage, from generation at the IoT level to processing, standardization, and exploitation at higher levels. Current DLT architectures for Industry 4.0 lack a holistic approach, as they often focus on securing data at its source (i.e., field level - IoT). Moreover, these architectures might not be completely suitable for the environments they are applied to, which leads to additional challenges.

Scalability and performance, for instance, are primary concerns [30]. Industry 4.0 systems generate enormous amounts of real-time data, and traditional unoptimized DLT platforms may struggle to handle such high transaction volumes and data throughput efficiently. This typically lead to performance bottlenecks and hinder the seamless integration of DLTs with existing industrial processes. Furthermore, the energy consumption and environmental impact of certain DLT implementations, especially those based on energy-intensive consensus mechanisms such as Proof of Work (PoW), might pose challenges for Industry 4.0 applications, which often prioritize sustainability and resource efficiency [31]. Thus, there is a clear need to develop lightweight DLT schemes.

Interoperability and compatibility also pose significant challenges not only in Industry 4.0, but also in DLTs [32]. Typically, systems involve various devices, protocols, and data formats, necessitating seamless communication between different DLT platforms and existing systems to achieve transparency, data sharing, and collaboration across the industrial ecosystem. At the core of this challenge is the fact that different DLT platforms usually employ unique protocols and data formats. Since DLTs are still an evolving technology, there is not yet a standardized approach that all platforms adhere to [33]. As a result, achieving interoperability between different DLTs can be a complex task, requiring the development of customized solutions or bridge technologies to facilitate communication between these disparate systems.

However, smart contracts come with inherent limitations that make accessing off-chain data directly, including standardized industrial data models, a challenge [34]. By their design, smart contracts live within the confines of their network, rendering them unable to directly interact with information outside their shared ledger. This poses a significant hurdle as the real-world, off-chain data is crucial in many industrial applications, particularly where adherence to standardized data models is imperative.

Data privacy and confidentiality are other aspects that cannot be overlooked. Busi-

nesses in Industry 4.0 often need to safeguard sensitive information such as trade secrets and intellectual property. Thus, DLTs must also implement specific mechanisms to ensure privacy and confidentiality [35].

Finally, the most relevant monetary costs associated with using DLTs largely revolve around the computational expenses for transaction validation [36]. These costs can be particularly significant when scaled to the high volume of transactions typical in industrial settings. Additional costs include transaction fees for network operations and smart contract execution, energy expenses for running nodes, and data storage costs for maintaining the ledger [37]. These financial commitments need to be carefully considered in the context of the long-term value proposition offered by DLTs.

A further more comprehensive analysis of this topic is presented in Chapter 3.

Summarizing, when implementing DLTs in Industry 4.0, **several requirements must be taken into account** as follows:

- The DLTs must cover the whole lifecycle of the data, from the moment it is generated from the machines until it is exploited for business logic. In order words, there is a need for a holistic approach.
- The DLTs need to be lightweight: efficient and scalable to process large amounts of data, and energy efficient.
- The DLTs must offer interoperability capabilities.
- The DLTs must have smart contracts, which also must have the capacity to interact with external environments.
- The DLTs should provide privacy, immutability and traceability.
- The DLTs must have low associated costs.

1.2 Thesis Main Objective and Research Questions

1.2.1 Objective

It is crucial to realize that the aforementioned obstacles are not insurmountable. As with any technological innovation, challenges are expected in early implementations and can be resolved through iterative refinement and adaptation [38]. The unique advantages of DLTs are compelling reasons to continue exploring and improving upon their implementation in Industry 4.0.

Therefore, the main **objective** of this thesis is:

> To study DLTs in depth, with a view to design a holistic DLT architecture that covers the whole cycle of the data (from when it is generated from the IoT machines up until it is exploited for business purposes) and addresses the aforementioned Industry 4.0 challenges without neglecting the particular challenges and requirements of the actual DLT technologies.

The design problem of the thesis can be defined using Wieringa's [39] template:

Improve *Industry 4.0 challenges hindering its adoption and development*

By *Designing an Industry 4.0 oriented architecture based on DLTs*

That satisfies *multiple criteria: covering the entire data process from generation to business logic exploitation; being efficient and scalable for processing large data volumes; offering energy efficiency; ensuring interoperability among different systems and devices; incorporating smart contracts with external environment interaction capabilities; providing data privacy, immutability, and traceability; and incurring low associated costs*

In order to *build an Industry 4.0 environment that has been able to overcome the challenges described above and thus facilitate its adoption*

1.2.2 Research Questions

The aforementioned problem statement, challenges, requirements, and objectives necessitate a granular and rigorous approach to dissect and address the complex nature of integrating DLTs into Industry 4.0. It is within this context that we introduce specific research questions, each designed to illuminate a facet of this multifaceted issue. The following research questions aim to provide actionable insights that contribute to achieving the main objective: **designing a comprehensive DLT architecture that suits the particular needs and challenges of Industry 4.0**. Thus, designing a DLT architecture for Industry 4.0 may first require revisiting certain aspects that are related to the pure nature of the current industrial standards.

> **Question 1. —** *Is the automation pyramid the ideal model on which we can design an architecture based on DLTs for Industry 4.0?*
>
> — This question is answered in Chapter 2, "Industry 4.0 Automation Pyramid: Revisited"

The first research question aims to examine the continued relevance of the traditional automation pyramid in the context of Industry 4.0, specifically when incorporating DLTs. This inquiry is critical because the automation pyramid has historically provided the architecture that underpins the flow of information in industrial settings, but Industry 4.0 introduces disruptive technologies such as DLTs that may challenge this hierarchical model. The question serves to probe whether this time-tested model is agile enough to adapt to the novel requirements of Industry 4.0, which extends beyond single-plant operations to encompass a complex web of multiple external actors and cutting-edge technologies. The answer to this question will inform the design principles for a DLT-based solutions optimized for Industry 4.0, making it foundational to the research.

Question 2. — *How can we design an architecture for Industry 4.0 based on DLTs that satisfies the above requirements?*

— This question is answered in Chapter 4, "A DLT-based Architecture for Industry 4.0"

The core endeavor of this thesis is to embark on a rigorous exploration aimed at developing a DLT-based architecture tailored for Industry 4.0. This architecture seeks to address an array of critical requirements, from the comprehensive handling of data as it moves from machine generation to business logic exploitation, to ensuring energy efficiency, interoperability and associated costs. By focusing on these specific parameters, the research aims to overcome the existing challenges that are impeding the broader adoption and development of Industry 4.0.

1.2.3 Contributions

Given the problem statement, challenges, objectives and research questions, this thesis aims at contributing on the following aspects:

1. First, a more suitable industrial scenario is elaborated upon, building on the existing pyramid of automation model. This is done to establish a robust and appropriate foundation for the requirements of Industry 4.0 and the incorporation of DLTs.

2. The current state-of-the-art in DLT architectures is studied, focusing particularly on their application in Industry 4.0, as well as open challenges and research opportunities.

3. Lastly, a DLT-based industrial architecture is designed to address existing challenges and facilitate the broader adoption of Industry 4.0. In this aspect further contributions have been made, primarily addressing the technical complexities encountered in each phase of the proposed architecture.

1.3 Methodology

This thesis has followed the Design Science Research (DSR) methodology [40]. DSR is commonly used in fields like information systems, computing, engineering, and other disciplines where the aim is not just to understand a problem but also to create innovative solutions for it. In contrast to purely empirical research approaches, which focus on observing and analyzing the world as it is, DSR aims to create new artifacts as a way of solving a problem and then to evaluate the effectiveness of those artifacts.

A.R. Hevner [41] proposes a three cycle process for DSR:

- **The Relevance Cycle** serves as the starting point for the DSR process by pinpointing and scrutinizing issues that need to be tackled within a specific context. The issue in question should be clearly defined and its significance within the context must be well-justified. The issue should also have broad appeal and any contributing factors to the problem may be identified and studied. Root-cause analysis can help systematically uncover the underlying reasons for the issue. While understanding the ramifications of a problem highlights its urgency, identifying its causes offers specific targets for rectification to prevent future occurrences. The Relevance Cycle not only identifies which issues should be addressed, but also sets the criteria for evaluating the effectiveness of the eventual solution.

- **The Design Cycle** serves as the central mechanism in any design science research initiative. It operates through a recurrent sequence of two key steps: initially, the creation and implementation of an artifact designed to resolve the problem as outlined in "The Relevance Cycle"; and subsequently, the assessment of the artifact's effectiveness. Insights gained from these evaluations can lead to further refinements of the artifact.

- **The Rigor Cycle** links the activities of design science research to an established foundation of scientific knowledge, field experience, and specialized expertise. This cycle enriches the research endeavor by incorporating past insights, thereby ensuring its innovative nature. It falls upon the researchers to rigorously explore and cite this wealth of knowledge to ensure that the outcomes of their work tran-

scend mere routine designs and contribute to research, rather than just applying well-understood methodologies.

This dissertation has been developed along the DSR hallmarks.

- A for the Relevance Cycle, in the context of overcoming the present challenges in Industry 4.0 adoption and development, two major problems that are particularly obstructive have been identified. First, there is a lack of a comprehensive, efficient, and scalable data architecture. Second, existing solutions often lack features such as energy efficiency, interoperability, smart contracts with external data access capability, and data security and privacy features. To comprehend these issues more deeply, a root-cause analysis has been carried out. Through this analysis, both the causal factors and potential consequences if these issues remain unresolved were examined.

  For each identified problem, a solution must fulfill some specific requirements. These include: A complete data lifecycle coverage, from data generation to business-logic exploitation; Efficiency and scalability in data processing; Energy efficiency; Interoperability with various systems and technologies; The inclusion of smart contracts with external interaction capability; Data privacy, immutability, and traceability features; and Cost-effectiveness.

- As for the Design Cycle, this research focuses on designing an Industry 4.0-oriented architecture based on DLTs. The aim is to satisfy all the requirements listed in the Relevance Cycle. This architecture will then undergo a series of evaluations to gauge its effectiveness in solving the identified problems.

- As for the Rigor Cycle, this project is deeply grounded in a rich body of scientific foundations, industry experience, and expertise. By extensively researching and referencing past works and existing technologies, it is ensured that the proposed architecture is not just a routine design but a significant research contribution. This also confirms that the architecture is innovative in addressing the unique challenges posed by Industry 4.0 adoption and development.

The overarching aim is to construct an Industry 4.0 environment that has effectively mitigated the challenges identified in the Relevance Cycle, thereby facilitating wider adoption and further development of Industry 4.0 technologies.

By aligning our research with the DSR cycles, a methodical, well-reasoned approach to solving critical issues in the advancement and adoption of Industry 4.0 is ensured.

1.4 Thesis Outline

This thesis showcases the outcomes of the efforts made to address the aforementioned research questions. A brief overview of the dissertation's structure and the key contributions of each chapter are provided below.

1.4.1 Chapter 2 - Industry 4.0 Automation Pyramid: Revisited

Chapter 2 presents the industrial case scenario on which is thesis is developed. The aforementioned scenario definition is built upon the automation pyramid model, a well-established structure in the manufacturing sector. This scenario consists of four tiers that represent different levels of industrial operations.

1.4.2 Chapter 3 - State of the Art

Chapter 3 presents a comprehensive Systematic Literature Review (SLR) of blockchain and other DLT architectures, specifically tailored for IoT-based fields (a subject of paramount significance in contemporary research), with a more specific focus on the scope of this thesis, Industry 4.0.

The SLR serves a critical function: it identifies gaps and underscores the imperative for a more integrative, holistic approach. This revelation becomes a muse, guiding the thesis project towards building a sweeping DLT-based solution. Such a solution is envisioned to not only counteract the frailties of existing frameworks but also venture into territories less charted, embracing the entirety of the data lifecycle, transcending the bounds of mere IoT consideration, and converging towards the multifaceted requirements of Industry 4.0.

1.4.3 Chapter 4 - A DLT-based Architecture for Industry 4.0

Chapter 4 presents the core contributions of this thesis. It introduces a multi-layered DLT architecture devised to enhance data management and security within an Industry 4.0 scenario. The architecture is structured across three distinct layers that are systematically arranged across the industrial process, from machine operation to high-level business decisions.

The presented architecture creates a comprehensive, secure, and efficient data handling system that runs through the entire industrial ecosystem. This integration enables seamless communication and information flow, leading to a more connected manufacturing environment.

1.4.4 Chapter 5 - Concluding remarks

This final chapter brings this thesis to a close by offering a summary of the key findings derived from the research conducted throughout this work. Additionally, this chapter presents the research outcomes in the shape of contributions published in specialized journals and presented at international conferences. Lastly, it highlights potential future research directions that have emerged as a result of the progress made in this thesis.

1.5 Summary and Conclusion

This chapter provided an introduction and comprehensive overview of the entire thesis. It set the foundation for the entire thesis by providing context, establishing the research questions and methodology, and outlining the structure of the work. Table 1.1 summarizes the aspects that have been presented in this chapter.

Problem definition Root Cause Analysis	**Design Problem (Wieringa's template)**	**Research Questions**
Context: Industry 4.0 **Problem:** The revolutionary promise of Industry 4.0 remains largely unrealized due to significant barriers to its adoption **Causes:** Cybersecurity, performance and scalability, data standardization and interoperability, automation and costs challenges **Consequences:** Loss of productivity and competitiveness, skilled labor, and customer satisfaction	**Improve** *Industry 4.0 challenges hindering its adoption and development* **By** *Designing an Industry 4.0 oriented architecture based on DLTs* **That satisfies** *multiple criteria: covering the entire data process from generation to business logic exploitation; being efficient and scalable for processing large data volumes; offering energy efficiency; ensuring interoperability among different systems and devices; incorporating smart contracts with external environment interaction capabilities; providing data privacy, immutability, and traceability; and incurring low associated costs* **In order to** *build an Industry 4.0 environment that has been able to overcome the challenges described above and thus facilitate its adoption*	**RQ1:** Is the automation pyramid the ideal model on which we can design an architecture based on DLTs for Industry 4.0? **RQ2:** How can we design an architecture for Industry 4.0 based on DLTs that satisfies the above requirements?

Table 1.1 Introduction summary

Chapter 2

Industry 4.0 Automation Pyramid: Revisited

2.1 Overview

This chapter provides the contextual foundation for the thesis within Industry 4.0, setting the stage for the subsequent detailed exploration of current advancements in blockchain and other DLTs, as well as the conception and design of an encompassing DLT architecture for Industry 4.0, which will fulfill the main objective of this thesis. The context is illustrated through an applicable industrial case scenario that exemplifies the main layers of the Industry 4.0 framework. This scenario is broken down into four tiers, inspired by the automation pyramid scheme, and concentrates on the optimization of assets, production lines, industrial facilities, and the ultimate unification of multiple plants into an Industry 4.0 business consortium. The objective of this approach is to address the critical pillars of Industry 4.0, including efficiency, security, interoperability, and traceability, with the ultimate goal of constructing a comprehensive DLT solution atop this framework.

Section 2.2 introduces the current industrial pyramid based framework and analyses its limitations before proposing a new adaptation for this thesis. Sections 2.3, 2.4, 2.5 and 2.6 gather the proposed levels of the aforementioned adaptation of the current industrial framework. Finally, Section 2.7 presents the summary and conclusions of the chapter.

2.2 Introduction

As previously touched upon in the introduction, Industry 4.0 heralds an era where digital innovations and data-driven insights are redefining industrial landscapes, promising unparalleled boosts in efficiency, productivity, and competitiveness [42].

Delving deeper into the architectural backbone of modern factories, a structured hierarchy, seamlessly connecting the hands-on tasks of the shop floor with the strategic intricacies of the ERP systems can be observed. This multilayered structure is elegantly captured by the automation pyramid, which, as introduced earlier, serves as a cornerstone reference for industrial initiatives. More than a mere depiction of hierarchical control, the pyramid elucidates the synergy of technology integrations, dynamic data exchanges, and diverse communication protocols across manufacturing spectrums, as depicted in Figure 2.1.

Furthermore, the role of the pyramid actively guides industries in the transition towards a more interconnected, agile, and adaptive operational approach. This perspective ensures that businesses remain resilient and responsive to the ever-evolving challenges of the contemporary industrial era. Embracing this model grants industries a clear roadmap, aiding in the intricate orchestration of Industry 4.0 components. A profound grasp of each tier of this pyramid equips industries with the insights to design and deploy transformative strategies, amplifying their operational quality and ensuring a forward-thinking approach in an increasingly competitive market [43].

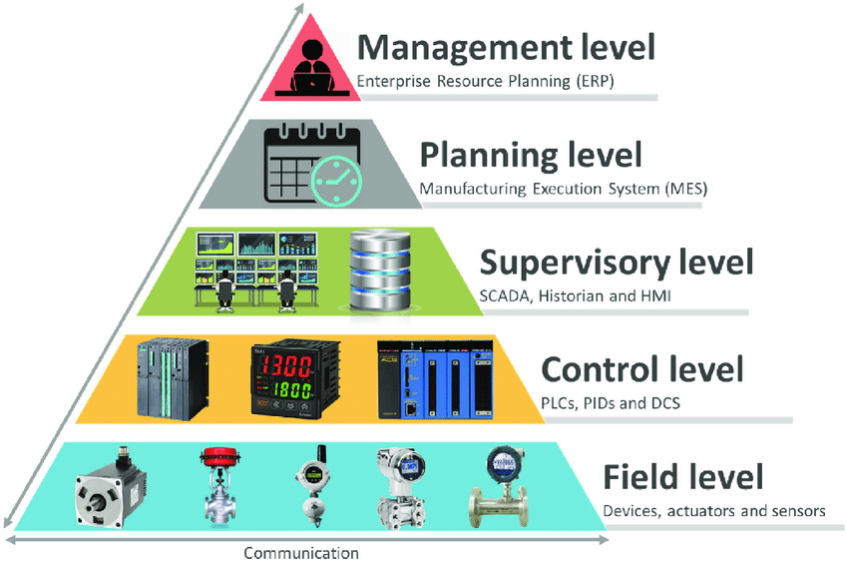

Fig. 2.1 Automation pyramid diagram [4]

However, while the state-of-the-art has seen some efforts to modernize the traditional pyramid of automation, these adaptations primarily focus on integrating specific Industry 4.0 technologies such as AI and Big Data [44]. Despite these updates, existing models still present several limitations that could obstruct the successful implementa-

tion of a DLT-based architecture in the context of Industry 4.0. These shortcomings are particularly significant given the challenges outlined in the introduction of this thesis.

Firstly, the pyramid structure is constrained by its scope, focused mainly on components within a single plant or facility. This narrow focus could hinder the development of architectures that aim to incorporate more expansive, interconnected systems. Industry 4.0, on the other hand, envisions a world where devices, systems, and even entire factories are interconnected, sharing data and making decentralized decisions. With the advent of DLTs, the possibility to ensure secure, transparent, and immutable records of transactions between interconnected devices and systems can be fully realized [45]. To accommodate this complexity, it might be necessary to move beyond the traditional pyramid model to a more flexible and inclusive model.

Secondly, the automation pyramid does not naturally account for external actors, which are increasingly relevant in the Industry 4.0 context [46]. Whether there are suppliers contributing to a seamless supply chain, or customer-side analytics that feed into production decisions, these external factors are becoming integral to modern manufacturing paradigms. The need for real-time data sharing and complex transactional interactions between multiple parties can be best facilitated through DLTs, which excel at providing a single source of truth in a decentralized system [47].

Therefore, starting from the aforementioned model, an industrial case scenario that serves as the basis for the work presented in this thesis is derived. The scenario consists of four levels, as illustrated in Figure 2.2, which is adapted from the original automation pyramid scheme. However, slight modifications have been made to better align with real-world industrial experiences and requirements. The four-layer industrial scenario is based on the automation pyramid since it serves as a well-established and widely-accepted reference model for the hierarchical structure of automation and control systems in the industrial domain [48]. The automation pyramid already offers a systematic approach to understanding the different layers of technology, data flow, and communication protocols involved in manufacturing processes. By leveraging the structure of the automation pyramid, a comprehensive and scalable solution that addresses the various aspects and challenges of implementing Industry 4.0 can be developed.

Adding to this, the adapted automation pyramid in this study benefits from the extensive industrial experience accrued at Ikerlan[1]. The model has been subjected to rigorous testing and validation in diverse industrial settings, thus adding an additional layer of credibility and practical relevance to the academic exploration of Industry 4.0 concepts. Not only does this bridge the gap between theory and application, but it also helps in fine-tuning the model to better represent the intricacies and nuances encoun-

[1] `https://ikerlan.es`

tered in actual industrial processes.

Since Ikerlan has a long-standing history of engaging with the industry, the model integrates industry-specific challenges and requirements, thereby making it more adequate for implementation and evaluation in future industrial case studies. These industry-informed modifications to the automation pyramid model allow for an enhanced understanding of the scalability and interoperability issues that are pivotal in the transition towards Industry 4.0. This ensures that the model is not just theoretically sound but also practically feasible, reinforcing its utility as a reliable framework for both researchers and industry practitioners alike.

The lessons learned from these real-world applications have fed back into the academic research, creating a virtuous cycle that enhances both the theoretical framework and its practical applications.

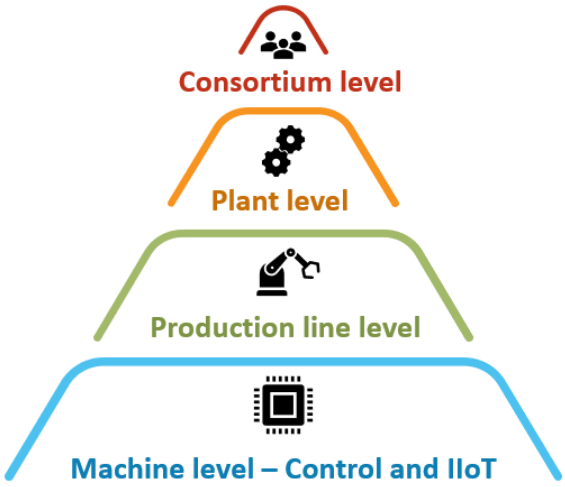

Fig. 2.2 Industry 4.0 case scenario

The levels of the automation pyramid have been adapted to this thesis' model in the following manner:

1. Machine level (Field and Control Levels): Combining the field and control levels of the automation pyramid, the focus is on optimizing individual assets, such as machines, sensors, actuators, and other components. This involves enhancing the performance and efficiency of each asset, as well as ensuring seamless integration with other components and control systems.

2. Production Line level (Supervisory Level): Drawing from the supervisory level of the automation pyramid, the second layer involves the optimization of entire production lines. This includes coordinating the operation of multiple assets, mini-

mizing bottlenecks and downtime, and ensuring optimal throughput and product quality.

3. Plant level (Manufacturing Operations Management Level): Mirroring the Manufacturing Execution System (MES) level of the automation pyramid, the third layer in this model focuses on optimizing the performance of entire industrial plants. This encompasses managing the production lines data, resources, energy consumption, waste reduction, and other aspects that contribute to the overall efficiency and sustainability of the plant.

4. Consortium level (Enterprise Resource Planning Level): Inspired by the ERP level of the automation pyramid, the fourth and top layer explores the integration of multiple industrial plants into an Industry 4.0 business consortium. This allows for greater collaboration, data sharing, and resource optimization among various stakeholders within the industrial ecosystem.

The defined case scenario serves as a foundation for designing the core contribution of this thesis, which is a holistic DLT architecture that addresses key aspects of Industry 4.0, including efficiency, security, interoperability, and traceability, throughout the whole lifecycle of the data, from when it is generated at the machine level and processed at the plant level, to when it is exploited at the consortium level for aggregated value. Thus, by systematically addressing each level of the pyramid, a comprehensive and scalable solution that is capable of meeting the diverse needs and challenges of the manufacturing sector in the context of Industry 4.0 can be developed.

2.3 Machine level

The machine level, which corresponds to cyber-physical production systems, is a crucial component of Industry 4.0 and the Industrial IoT (IIoT). This level is responsible for generating a significant amount of data, which necessitates efficient and real-time management to ensure optimal performance and decision-making in industrial processes [49]. The sheer volume of critical data generated at this level not only affects the security and privacy of the underlying systems but also has implications for the humans interacting with these systems [50]. In addition, many lightweight devices operating at this level are powered by batteries, making energy efficiency a crucial challenge and requirement for the sustainability and longevity of these devices and the overall system [24].

The machine level is composed of several types of devices that play distinct roles in the overall functioning and management of industrial processes. These devices can be

classified into the following categories:

- Sensors: IIoT devices specifically designed to measure various system parameters, such as temperature, pressure, humidity, flow rate, and vibration, among others. Sensors collect data from the physical environment and convert it into digital signals that can be processed, analyzed, and acted upon by other components in the system.

- Actuators: Actuators are IIoT devices responsible for executing specific actions in response to input from sensors or control devices. These actions can include opening or closing valves, starting or stopping motors, adjusting equipment settings, and more. Actuators are essential for implementing closed-loop control systems that automatically adjust and optimize industrial processes based on real-time data.

- Control devices: These devices and systems, such as PLC, RTU, and Distributed Control System (DCS), control industrial processes based on information received from IIoT devices like sensors and actuators [51]. Control devices are responsible for processing sensor data, making decisions, and sending commands to actuators to adjust the operation of machines and equipment in accordance with the desired outcomes. Human-Machine Interface (HMI) and Supervisory Control And Data Acquisition (SCADA) systems are also part of the control devices category, enabling operators to monitor and interact with the industrial processes [51]. Control systems are designed to manage multiple IIoT devices, ensuring that the entire process runs smoothly and efficiently [52].

- Predictive Maintenance Systems: Predictive maintenance devices and systems use advanced analytics and machine learning algorithms to analyze data from sensors and other sources to predict potential equipment failures and suggest maintenance actions before a failure occurs. These systems help minimize downtime, reduce maintenance costs, and extend the life of equipment.

The integration and interaction of these devices at the machine level form the foundation for advanced cyber-physical production systems. By ensuring efficient data generation, management, and decision-making, as well as addressing the challenges of security, privacy, and energy efficiency, the machine level plays a critical role in realizing the full potential of Industry 4.0 and the IIoT.

2.4 Production line level

Production lines are essential components of contemporary manufacturing processes, consisting of numerous interconnected machines, robots, and other equipment that work together in a coordinated and efficient way [53]. In the context of a smart factory, individual production lines form sub-networks that contribute to the overall performance and productivity of the facility. The challenges and demands at this level are akin to those of the machine level, but with unique aspects due to the increased intricacy and inter-connectivity of the various components involved.

Rapid and secure handling of vast amounts of data is a critical requirement for modern production lines. As machines and robots generate data, the production lines must efficiently process, analyze, and respond in real-time to maintain peak performance, avoid errors, and minimize downtime. Data security and privacy must also be prioritized, as unauthorized access or tampering might lead to substantial disruptions, safety hazards, or loss of valuable intellectual property.

Another crucial concern for production lines is reliability, as inconsistent boot-up behaviors and extended start-up times can negatively impact the efficiency and output of the manufacturing process. Addressing these issues through solid system design, testing, and continuous monitoring helps ensure that equipment and processes operate reliably and consistently, reducing the risk of downtime and production losses.

Scalability is vital for production lines, given the large number of IIoT devices and the significant data traffic generated. This allows for adjustments in production demands and future expansion. By designing and implementing adaptable and modular systems that can be easily modified, upgraded, or scaled up as necessary, organizations can adapt to shifting market conditions, embrace emerging technologies, and maintain a competitive edge.

Energy and cost efficiency are also critical considerations at the production line level. With increasing energy costs and stricter environmental regulations, manufacturers must optimize energy consumption while preserving productivity and product quality. The incorporation of energy-efficient technologies, such as variable frequency drives, energy recovery systems, and smart lighting, can lead to reduced energy consumption and lower operational costs. Moreover, adopting lean manufacturing principles, process automation, and predictive maintenance strategies can help minimize waste, streamline operations, and enhance cost efficiency further.

In summary, modern production lines play a pivotal role in the overall performance and success of smart factories within the Industry 4.0 paradigm. By addressing the challenges and requirements associated with data management, security, reliability, scalabil-

ity, energy efficiency, and cost efficiency, organizations can develop flexible, resilient, and high-performing production lines that deliver exceptional value and drive growth in an increasingly competitive and rapidly evolving industrial environment [54].

2.5 Plant level

The plant level encompasses all production lines within a smart factory, integrating them into a single, comprehensive network that enables seamless communication and coordination across the entire facility [55]. This level presents numerous challenges, both in terms of physical constraints and operational requirements, which must be addressed to ensure the optimal functioning of the smart factory as a whole.

One significant challenge at the plant level is the size of the network and the distances between its component nodes. This is particularly relevant for wireless connections, where latency can become a critical concern due to the physical separation between devices and the need for real-time communication and data processing [56]. To mitigate latency issues, advanced networking technologies, such as 5G, can be implemented to provide high-speed, low-latency connections that support the stringent requirements of modern smart factories.

Another important consideration at the plant level is the need for a highly optimized, secure, and efficient distributed storage solution to manage the vast amount of data generated and processed within the factory [24]. This requires the implementation of robust storage systems, such as distributed databases or edge computing infrastructures, that can handle high volumes of data, ensure data integrity and security, and provide rapid access to information when needed. Additionally, advanced data analytics and machine learning techniques can be employed to derive valuable insights from the data, facilitating improved decision-making, process optimization, and predictive maintenance.

Interoperability is a key requirement at the plant level, as it enables seamless collaboration between various industrial machines and plants that may have diverse and modular components [54] [57]. This necessitates the adoption of standardized communication protocols, data formats, and system architectures that ensure compatibility between different devices, systems, and software applications. Moreover, implementing a flexible and modular system design allows for the integration of new technologies and the adaptation to evolving industry standards, ensuring that the smart factory remains agile and future-proof.

To further enhance the plant level, it is also essential to consider the role of human operators, who are required to interact with and oversee the various processes within the smart factory. Providing user-friendly interfaces, comprehensive training programs, and

clear procedures can help to ensure that operators can effectively manage the complex, interconnected systems at the plant level, enhancing overall productivity and safety.

In summary, the plant level plays a pivotal role in the overall functioning and success of smart factories in the Industry 4.0 era. By addressing the challenges associated with network size and latency, data storage, interoperability, and human-machine interaction, organizations can create a highly integrated, efficient, and adaptable plant-level infrastructure that supports the seamless operation and growth of the smart factory in an increasingly competitive and rapidly evolving industrial landscape.

2.6 Consortium level

The consortium level represents the highest level of integration in the Industry 4.0 paradigm, bringing together multiple smart factories into a consortium network that facilitates governance, interoperability, traceability, and secure, private distributed storage for the entire ecosystem of smart factories [58]. Achieving effective coordination and collaboration at this level requires addressing several challenges and ensuring that the network meets the unique needs of a diverse and distributed set of participants.

One notable challenge at the consortium level is the physical separation of the nodes, which can span hundreds or even thousands of kilometers. This necessitates the implementation of advanced networking technologies that can provide reliable, high-speed communication across vast distances. Technologies such as dedicated fiber-optic lines, satellite communication, and next-generation wireless networks (e.g., 5G) can be employed to facilitate seamless communication and data exchange between smart factories within the consortium.

High transaction speed is another essential requirement at the consortium level, as the network must support rapid and efficient data exchange, decision-making, and coordination between multiple organizations. Optimized data processing and communication methods, such as edge computing, parallel processing, and real-time data analytics, can help to ensure that consortium members have access to the information they need when they need it, enabling them to make timely and informed decisions.

Transparency and immutability of data are critical to fostering trust and accountability within the consortium. This can be achieved through the use of blockchain technology, which provides a decentralized, tamper-proof ledger that records all transactions and data exchanges within the network. The blockchain's inherent transparency and immutability serves to ensure that all consortium members have access to accurate, up-to-date information and can verify the integrity of the data at any time.

Smart contracts and financial transactions are additional capabilities that the con-

sortium network should support, as they enable automated, secure, and efficient interactions between multiple organizations. By leveraging blockchain technology and integrating smart contract functionality, the consortium network can facilitate the execution of complex, multi-party agreements and transactions, streamlining business processes and reducing the need for manual intervention and third-party intermediaries.

Moreover, the consortium network should prioritize security and privacy to protect sensitive data and intellectual property while enabling collaboration between members. This can be achieved through the implementation of robust encryption, access control, and data anonymization techniques, as well as the establishment of clear policies and procedures governing data sharing and usage within the consortium.

In conclusion, the consortium level plays a critical role in enabling seamless cooperation and coordination between multiple smart factories within the Industry 4.0 landscape. By addressing challenges related to network connectivity, transaction speed, data transparency and immutability, smart contracts and financial transactions, and security and privacy, organizations are able to create a highly efficient, secure, and collaborative consortium network that drives innovation, enables trust, and accelerates the adoption and growth of Industry 4.0 technologies across the entire industrial ecosystem [59].

2.7 Summary and Conclusion

This chapter provided a rich examination of the transformative impact of Industry 4.0 on the manufacturing sector. The shift towards this industrial revolution signifies a leap towards enhancing the efficiency, productivity, and competitiveness of industrial plants. This is achieved through the integration of cutting-edge digital technologies, data-driven insights, and interconnected systems.

The automation pyramid, representing the hierarchical structure of modern factories, forms the cornerstone of this industrial transformation. The model establishes a systematic methodology for understanding and implementing the diverse layers of technology, data flow, and communication protocols involved in the manufacturing process. This tool enables businesses to navigate the transition to Industry 4.0 successfully.

Significantly, this chapter presented an adaptation of the automation pyramid into a four-layer industrial scenario that is used as the starting context of this work. This adaptation elucidates a structured approach to adopting Industry 4.0, with each layer — machine, production line, plant, and consortium — addressing distinct facets of the industrial process. From optimizing individual assets and entire production lines to the integration of multiple plants into a business consortium, this model provides an overarching strategy for the application of Industry 4.0 principles.

Chapter 3

State of the Art

3.1 Overview

With the increasing integration of DLTs -being blockchain the most common DLT type- within IoT ecosystems, there is a clear need for strategies that fit the resource constraints typical of IoT devices. Traditional blockchain approaches, though promising in terms of decentralization, security, and privacy, often falter in these settings due to high resource consumption and limited throughput. Therefore, the emphasis has largely shifted towards the conceptualization and development of lightweight DLT solutions compatible with IoT environments. This movement within the academic and technological communities necessitates a structured and in-depth examination to guide future DLT architecture implementations.

While the core of DLT research within the IoT sphere is undeniably significant, it is imperative to examine this integration within a broader context, spanning diverse domains such as Industry 4.0, smart cities, smart homes, etc. Given Industry 4.0 emphasis on the digital transformation of manufacturing and industrial operations, a holistic approach that takes into account the entire industrial ecosystem is warranted. This includes IoT devices, interconnected plant systems, and diverse stakeholders.

In light of this, this state-of-the-art study adopts a dual-fold approach: first, to offer a comprehensive review of the current state of DLT architectures, with a keen eye on performance limitations and innovations in areas with a strong IoT presence, especially Industry 4.0; and second, to pinpoint existing research gaps, offering a roadmap for a more unified and robust DLT architecture tailored for Industry 4.0. The overarching objective is to derive insights from this expansive landscape, address the extant limitations of current architectures, and harness unexplored research trajectories. In doing so, this study endeavors to lay the groundwork for a cohesive DLT blueprint adeptly suited to the multifarious challenges and prerequisites of Industry 4.0.

Section 3.2 presents the introduction of this study. Section 3.3 presents the related surveys on this topic. Section 3.4 describes the studied research questions and research methodology. Section 3.5 presents the results of the study attending the defined research questions. Section 3.6 discusses the results of the study. Finally, Section 3.7 presents the summary and conclusions of the chapter.

3.2 Introduction

The incorporation of DLTs -mostly blockchain- within IoT ecosystems has been increasingly explored as a potential solution to address the current limitations and challenges faced by centralized architectures. The promise of decentralized control, robust security, and enhanced privacy are some of the core benefits offered by DLTs. However, the application of DLTs, especially blockchain, in resource-constrained environments with IoT devices presents unique challenges, as traditional consensus algorithms often exhibit high resource consumption, limited throughput, and transaction delay [60]. Moreover, the diverse requirements of various IoT use cases necessitate tailored and efficient solutions [61]. Consequently, significant scientific efforts have been directed towards creating lightweight DLT architectures. This necessitates a comprehensive systematic analysis of the current state-of-the-art in order to inform future optimizations and novel architectures for various applications.

While the IoT ecosystem itself is a prominent area for DLT exploration, it is essential to recognize the broader applicability of DLTs across various domains, including Industry 4.0, smart homes, smart cities, etc. Industry 4.0, in particular, is an emergent field focused on the digital transformation of manufacturing and industrial operations through the integration of advanced technologies such as IoT, AI, robotics, and data analytics. A holistic Industry 4.0 architecture necessitates consideration of the entire industrial ecosystem, encompassing not only IoT devices but also other interconnected systems and stakeholders. As such, exploring the adoption and adaptation of DLT technologies across diverse domains provides valuable insights and lessons for the design of a robust and scalable Industry 4.0 architecture.

This study aims to broaden the scope of investigation beyond the narrow confines of IoT-specific DLT solutions. Since this thesis is focused on the Industry 4.0 field, firstly a search for papers in this field was conducted using the following terms:

"Industry 4.0" AND "Blockchain" OR *"Distributed ledger"* OR *"DLT"*

However, the results of this search return a highly limited number of articles, many of which are survey papers rather than specific architectural proposals. Thus, given the

limited number of proposals specifically addressing Industry 4.0 and the inherent interconnections of IoT devices within industrial environments, it is crucial to analyze DLT applications across a range of fields, including smart homes, smart cities, and more. And since the DLT field is broad, there is a need to focus on the concept of "lightweight" DLTs, since according to the existing literature [62] [63] [64], most researchers are focusing of the design of "lightweight" solutions that can fulfill the requirements of IoT environments due to the aforementioned performance issues of DLTs. Hence, this approach allows for the identification of cross-domain synergies and the extrapolation of best practices that can be adapted to the requirements of Industry 4.0. The Venn Diagram presented in Figure 3.1 shows the relation between DLT architectures that were specifically tailored for Industry 4.0 and overall architectures that were categorized as "lightweight".

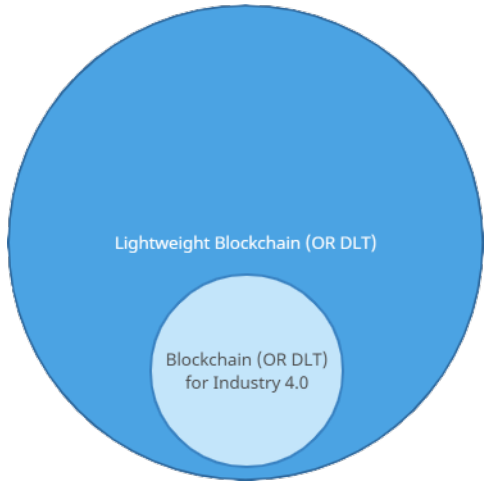

Fig. 3.1 Venn diagram of blockchain based architectures

This study assesses the architectural designs, consensus mechanisms, and other technical characteristics of DLT technologies across various domains. In general, this preliminary literature study of the thesis analyses the trade-offs involved in implementing DLTs in resource-constrained environments with IoT devices. The outcomes of this analysis serve as a serves as a preliminary basis for the future development of a holistic and adaptable DLT-based architecture for Industry 4.0.

3.3 Related studies

A total of 24 related reviews of DLTs for many environments have been identified. A Google Scholar search using the following string has been used:

"Blockchain" OR *"DLT"* AND *"survey"* OR *"review"* OR *"state of the art"*

The main contributions of each relevant survey are summarized below. Table 3.1 offers a comprehensive classification and comparison of the related work, while also emphasizing the focus and contributions of the present study.

Table 3.1 Related work comparison

Ref.	Systematic review	"Lightweight"	Evaluation	Technical aspects	Research opportunities
[65]					✓
[62]				✓	✓
[63]				✓	✓
[64]				✓	✓
[66]					
[67]	✓				
[68]				✓	
[69]				✓	✓
[70]					✓
[71]	✓		✓		
[72]	✓		✓		✓
[73]					
[22]					
[74]				✓	
[75]				✓	
[76]				✓	
[77]					
[78]					✓
[79]				✓	✓
[80]					✓
[81]			✓	✓	
[82]				✓	✓
[83]					✓
[84]		✓	✓		
This	✓	✓	✓	✓	✓

T. M. Fernández-Caramés and P. Fraga-Lamas [62] presented a thorough review on how to adapt blockchain to the specific needs of IoT in order to develop Blockchain for IoT (BIoT) applications. M. Wu *et al.* [63], M. S. Ali *et al.* [64], D. A. Noby and A. Khattab [66], F. A. Abadi *et al.* [67] and Y. Mezquita *et al.* [68] conducted comprehensive surveys on the applications of blockchain in IoT. H-N. Dai *et al.* [69] provided an overview of blockchain and its convergence with IoT by presenting a proposal of Blockchain of Things (BCoT). R. A. Memon *et al.* [70] provided a taxonomy of the

challenges in the current IoT infrastructure and a literature survey with a taxonomy of the issues to expect in the future of IoT after adopting blockchain. M. Conoscenti *et al.* [71] tried to understand whether the blockchain and Peer-to-Peer (P2P) approaches can be employed to foster a decentralised and private-by-design IoT. S. K. Lo *et al.* [72] focused on analysing the solutions proposed in academia and the methodologies used to integrate blockchain with IoT. Q. Wang *et al.* [73] and T. Alladi *et al.* [22] discussed the integration of blockchain and IoT but only for one specific application: the IIoT. L. Lao *et al.* [74] analysed popular blockchain-IoT architectures but only discussed their consensus algorithms. Finally, B. Farahani [75] presented challenges, opportunities, applications and solutions of blockchain for e-health. X. Wang *et al.* [76] and P. Karthikeyyan *et al.* [77] surveyed the current limitations and security issues of IoT. J. Sengupta *et al.* [78] surveyed the attacks and security issues of blockchain when applied to IIoT. M. Khan and K. Salah [85] and M. Alamri *et al.* [79] discussed how blockchain could be a key enabler in solving many IoT security problems. M. A. Ferrag *et al.* [80] provided a classification of threat models considered by blockchain protocols in IoT networks and a taxonomy and a side-by-side comparison of the state-of-the-art methods towards secure and privacy-preserving blockchain. S. Madumidha *et al.* [81] and F. Lin *et al.* [82] focused on the applicability of blockchain for IoT in order to tackle security issues. M. Alizadeh *et al.* [83] surveyed the most common attacks that affect blockchain networks and the solutions to mitigate them, intending to assess how malicious these attacks are in IoT.

In this state-of-the-art study, the focus diverges from the previous works mentioned. A targeted approach is adopted to assess the technical attributes of a considerable number of peer-reviewed DLT architecture proposals for IoT, explicitly categorized as "lightweight", since most authors focus on designing solutions that improve the performance of the initial DLT solutions. Thus, the concept of lightweight, transitioning from a broad perspective (i.e., definitions) to a nuanced examination (i.e., consensus mechanisms, storage approaches, cryptographic techniques, evaluations) of individual proposals is dissected.

The central objective of this study is to accentuate the specific technical facets, requirements, challenges, and trends in DLT development for application areas demanding efficient solutions. To the present day, no existing systematic review concentrates solely on the technical aspects of DLT architectures tailored for resource-constrained environments. There exists a concise review paper by Hanggoro et al. [84], which offers a summary of eight solutions identified as "lightweight blockchain". However, this review falls short in delivering a comprehensive analysis of the summarized works.

This study delves into the architectural designs, consensus mechanisms, crypto-

graphic approaches, and scalability solutions across the spectrum of efficient DLT technologies for IoT-based environments. The technical insights gleaned from this review will serve as a robust foundation for developing a holistic DLT-based architecture for Industry 4.0, accommodating the entire industrial ecosystem.

3.4 Method

This section states the method that was used to conduct the study. The method includes the search methodology and the used sources, the research questions, the eligibility criteria and the data collection process.

3.4.1 Research questions

The research questions that this study addresses are as follows:

- **RQ1.** What are the most common fields of application for DLT architectures?
- **RQ2.** What characteristics do the proposals have?
- **RQ3.** Which aspects do the authors optimize in their architecture?
- **RQ4.** How are the architectures evaluated?

RQ1 is focused on gathering information about the fields of application of the current DLT solutions. This information will provide an insight on the main fields of application where DLT solutions are being applied and possible opportunities. Since this thesis is oriented towards the field of Industry 4.0, the focus is mostly put on examining this field of application over the rest.

RQ2 pretends to study the main characteristics of the studied architecture proposals in order to perform a comprehensive comparison. The characteristics that will be gathered are as follows:

- The type of the blockchain (or DLT)
- The structure of the framework
- The consensus protocol
- The type of storage

RQ3 is pointed on studying the parts of the reviewed proposals that are mostly optimized in order to see which aspects of the DLTs are getting the most and the least attention from the researchers. The possible optimized, thus lightweight aspects will be classified as follows:

- Consensus
- Storage
- Architecture
- Cryptography

RQ4 intends to study the evaluation of each proposal in order to gather information about the existing platforms and methods of testing / evaluation as well as insights into how to properly build and test blockchain (and other DLTs) architectures.

3.4.2 Paper inclusion criteria

The selected papers on the topic must achieve all of the four inclusion criteria in order to be eligible for this review. These criteria were defined in order to provide the most adequate papers that would help us provide an answer to all the research questions and achieve the objectives of this study. The criteria and the corresponding explanation is shown in Table 3.2.

Table 3.2 Paper inclusion criteria

No	Criteria	Explanation
1	The study must be an original research paper that introduces a novel DLT framework or improves an inefficient aspect the aforementioned technology.	Survey and review papers will be excluded since they do not include a specific architecture proposal.
2	The proposed solution must be evaluated.	This review is focused on studying DLTs in a practical manner rather than just theoretically. Hence, each proposal has to be evaluated.
3	The reviewed paper must be written in English.	English is the standard-universal language for scientific papers.

3.4.3 The search and the paper sources

This study was conducted by manually searching through six of the most relevant scientific search engines:

- dblp (`https://dblp.org/`)
- Google Scholar (`https://scholar.google.es/`)
- Web Of Science (`http://wos.fecyt.es/`)
- Scopus (`https://www.scopus.com/search/form.uri`)

- IEEE Xplore (`https://ieeexplore.ieee.org/`)
- ACM (`https://dl.acm.org/`)

The search string for searching involved two main concepts: lightweight AND DLTs. The term "blockchain" is used as a search term since currently, blockchains are by far the most used type of DLT. The complete search terms are as follows:

"Lightweight" AND "Blockchain" OR *"Distributed ledger"* OR *"DLT"*

The choice of search terms in a state-of-the-art analysis directly dictates the nature and relevance of the materials identified, ensuring alignment with the core research objectives. In this study, the search terms re employed based on the following reasons:

- Contextual demand: Many non-financial applications leveraging DLTs are constrained by factors like bandwidth, storage capacity, processing power, and energy consumption. Use cases such as IoT devices and edge computing highlight these constraints. Consequently, a "lightweight" DLT architecture, optimized for these constraints, is a pivotal requirement for these applications.
- Research orientation: The central objective revolves around pinpointing DLT solutions tailored for non-financial scenarios where resources are limited. The term "lightweight" inherently signifies efficiency, scalability, and minimized resource consumption, which are the crucial parameters for the study.
- Relevance and Precision: While "blockchain / DLT" ensures we are homing in on materials discussing blockchain technology and other DLTs, the term "lightweight" refines the search further. It ensures identifying architectures specifically crafted or adapted for resource-constrained environments, bridging the research objective and the pertinent literature.
- Synergy of terms: By integrating "lightweight" with "blockchain" or "DLT", the search string is strategically designed to spotlight literature that delves into DLT solutions optimized for environments with limited resources. This combination ensures a balanced approach, where we do not overlook broader DLT innovations while also emphasizing the lightweight characteristic.
- Reinforcement from prior works: Previous studies have underscored the importance of lightweight DLT solutions, especially in contexts such as IoT and edge computing. This not only supports the choice of search terms but also highlights the broader academic and industry acknowledgment of the importance of lightweight DLT solutions in non-financial domains.

- Scope of review: The choice of these terms was also influenced by the need for a comprehensive review. By focusing on lightweight DLT concepts, this works aims to encompass a broad spectrum of studies, ensuring that it encapsulate all relevant literature addressing the unique demands of non-financial environments.

In summary, the selection of "lightweight" DLTs as the primary search terms is rooted in the specific demands of this research context and objectives. This choice guarantees the identification of literature that is not only pertinent but also aligns with the unique requirements of blockchain applications in non-financial, resource-constrained settings.

3.4.4 Study protocol and process

This study was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [86], which diagram is shown in Figure 3.2. The PRISMA protocol was employed as it offers several key benefits:

- It demonstrates a quality review.
- It allows readers to assess strengths and weaknesses.
- It permits the replication of the reviewing process.
- Its structure is compatible with the standard guidelines for systematic literature reviews in computer science proposed in [87].

After retrieving the available articles following the defined research terms from the databases and removing the duplicates, each article's title and abstract were screened independently for eligibility using the criteria defined in Table 3.2. From a total of 251 papers, 115 were positively evaluated. Furthermore, 11 more papers were included in the study based on a reference follow-up of the papers that were initially elected, making for a total of 126 included papers. Table 3.3 shows the reference and titles of the included papers.

Table 3.3 Reviewed papers

Ref.	Title
[60]	A Fast Lightweight Consensus Algorithm for IoT Applications
[88]	Lightweight Blockchain to Improve Security and Privacy in Smarthome
[89]	A Lightweight Scalable Blockchain for IoT security and anonymity
[90]	An efficient Lightweight integrated Blockchain (ELIB) model for IoT security and privacy
[91]	One drone one block-based lightweight blockchain architecture for internet of drones

Continued on next page

Table 3.3 – continued from previous page

Ref.	Title
[92]	A blockchain-based AI-empowered contagious pandemic situation supervision scheme using IoT drones
[93]	A federated learning-based blockchain-embedded data accumulation scheme using drones for IoT
[94]	A lightweight blockchain based framework for underwater IoT
[95]	A blockchain-based normalized autonomous transaction settlement system for IoT-based e-commerce
[96]	Lightweight blockchain assisted secure routing of swarm UAS networking
[97]	CrowdLBM: A lightweight blockchain-based model for mobile crowdsensing in the IoT
[98]	Lightweight Blockchain for Healthcare
[99]	A novel blockchain framework for industrial IoT edge computing
[100]	SCC: Storage Compression Consensus for Blockchain in Lightweight IoT Network
[101]	Zerocalo - A lightweight blockchain based on DHT network
[102]	LightBC: A Lightweight Hash-Based Blockchain for the Secured IoT
[103]	LVChain: A lightweight and vote-based blockchain for access control in the IoT
[104]	BlockTrack-L: A lightweight blockchain-based provenance message tracking in IoT
[105]	A lightweight permission-based blockchain for IoT environments
[106]	Lightweight End-to-End Blockchain for IoT Applications
[107]	Lightweight fog based solution for privacy-preserving in IoT using blockchain
[108]	TCON - A lightweight Trust-dependent Consensus framework for blockchain
[109]	Securing IoT network using lightweight multi-fog blockchain model
[110]	PCBChain: Lightweight Reconfigurable Blockchain Primitives for Secure IoT Applications
[111]	LDC: A lightweight data consensus algorithm based on blockchain for IIoT for smart city applications
[112]	Microchain: A Hybrid Consensus Mechanism for Lightweight Distributed Ledger for IoT
[113]	Towards secure network computing services for lightweight clients using blockchain
[114]	EPBC: Efficient public blockchain client for lightweight users
[115]	LightChain: On the lightweight blockchain for the IoT
[116]	Blockchain-based lightweight authentication mechanism for vehicular fog infrastructure
[117]	A lightweight blockchain based two factor authentication mechanism for LoRaWAN join procedure
[118]	Polynomial-based Lightweight Key Management in a Permissioned Blockchain
[119]	A Novel Enhanced Lightweight Node for Blockchain
[120]	Leveraging lightweight blockchain to establish data integrity for surveillance cameras
[121]	A Lightweight Blockchain-Based Model for Data Quality Assessment in crowdsensing
[122]	Blockchain-based lightweight trust management in mobile ad-hoc networks
[123]	PoBT: A Lightweight Consensus Algorithm for Scalable IoT Business Blockchain
[124]	Sensor-chain: A lightweight scalable blockchain framework for IoT
[23]	Lightchain: A lightweight blockchain system for IIoT
[125]	EC-ElGamal and Genetic Algorithm-Based Enhancement for Lightweight Scalable Blockchain in IoT
[126]	A lightweight hash-based blockchain architecture for IIoT
[127]	Mobile charger billing system using lightweight Blockchain
[128]	LayerChain: A Hierarchical Edge-Cloud Blockchain for Large-Scale Low-Delay IIoT Applications
[129]	LDV: A Lightweight DAG-Based Blockchain for Vehicular Social Networks
[130]	PoEWAL: A lightweight consensus mechanism for blockchain in IoT
[131]	A Lightweight Blockchain-based Technique for Anti-Tampering in Wireless Sensor Networks
[132]	Blockchain-based IoT access control system: Towards security, lightweight, and cross-domain

Continued on next page

Table 3.3 – continued from previous page

Ref.	Title
[133]	A lightweight scheme with dual-blockchain for intelligent pricing system of smart grid
[134]	Tikiri—Towards a lightweight blockchain for IoT
[135]	Lightweight Blockchain Framework using Enhanced Master-Slave Blockchain Paradigm
[136]	Lightweight and Scalable DAG based distributed ledger for verifying IoT data integrity
[137]	AEchain: A lightweight blockchain for IoT applications
[138]	Fusion chain: A decentralized lightweight blockchain for IoT security and privacy
[139]	Blockchain-Based Decentralized Lightweight Control Access Scheme for Smart Grids
[140]	Edge Blockchain Assisted Lightweight Privacy-preserving Data Aggregation for Smart Grid
[141]	Lightweight blockchain consensus mechanism and storage optimization for resource-constrained IoT
[142]	A scalable blockchain framework for secure transactions in IoT
[143]	DAAC: Digital Asset Access Control in a Unified Blockchain Based E-Health System
[144]	A scalable IoT protocol via an efficient DAG-based distributed ledger consensus
[145]	An Efficient and Compacted DAG-Based Blockchain Protocol for IIoT
[146]	DAG-based distributed ledger for low-latency smart grid network
[147]	A blockchain and IoT based framework for information transparency in supply chain finance
[148]	A blockchain-based Fog-oriented framework for smart public vehicular transportation systems
[149]	Accident responsibility identification model for Internet of Vehicles based on lightweight blockchain
[150]	AirBC: A Lightweight Reputation-based Blockchain Scheme for Resource-constrained UANET
[151]	A lightweight framework for secure virtual machine migration in cloud federations using blockchain
[152]	Design and Evaluation of a Heterogeneous Lightweight Blockchain-Based Marketplace
[153]	Design of Blockchain-Based Lightweight V2I Handover Authentication Protocol for VANET
[154]	LBlockchainE: A Lightweight Blockchain for Edge IoT-Enabled Maritime Transportation Systems
[155]	Lightweight Blockchain Security Protocol for Secure Storage and Communication in SDN-Enabled IoV
[156]	Blockchain Remote Mutual Authentication for AI-Empowered IoT Sustainable Computing Systems
[157]	Lightweight Blockchain Framework For Medical Record Data Integrity
[158]	Lightweight blockchain system for resource-constrained IoT devices
[159]	Lightweight Blockchain-Based Scheme to Secure Wireless M2M Area Networks
[160]	Lightweight branched blockchain security framework for Internet of Vehicles
[161]	Lightweight Modified Consensus Approach in IoT Blockchain
[162]	Blockchain for Secured Wireless Sensor Networks: Energy Consumption of MAC Address-Based PoA
[163]	Securing IoT Environment using Lightweight Blockchain Approach
[164]	A Lightweight Blockchain Framework for IoT Integration in Smart Cities
[165]	A lightweight blockchain-based access control scheme for the IoT
[166]	A Lightweight Smart Meter Framework using a Scalable Blockchain for Smart Cities
[167]	A two-layer blockchain mechanism for reliable crossing-domain communication in smart cities
[168]	A Study on Lightweight And Secure Edge Computing Based Blockchain
[169]	A Traditional Chinese Medicine Traceability System Based on Lightweight Blockchain
[170]	BAASH: Lightweight, Efficient, and Reliable Blockchain-as-a-Service for HPC Systems
[171]	Cybertwin Blockchain for Lightweight and Privacy-Preserving Authentication in Internet of Vehicles
[172]	Hardware/software co-design for lightweight blockchain-secured on-device machine learning
[173]	ECLB: Edge-Computing-Based Lightweight Blockchain Framework for Mobile Systems
[174]	LightBC: A Lightweight Hash-Based Blockchain for the Secured IoT

Continued on next page

Table 3.3 – continued from previous page

Ref.	Title
[175]	Lightweight and Reliable Decentralized Reward System using Blockchain
[176]	Lightweight Blockchain Based on Storage Resource Optimization for Internet of Vehicles
[177]	Lightweight blockchain consensus mechanism and storage optimization for IoT devices
[178]	Lightweight Blockchain Secured Framework for Smart Precise Farming System
[179]	Lightweight blockchain to solve forgery and privacy issues of vehicle image data
[180]	A Lightweight Security Mechanism for IoT Based Smart City Management Systems using Blockchain
[181]	Secure and Privacy-Preserving Lightweight Blockchain for Energy Trading
[182]	A Lightweight Architecture Based on DAG-Lattice Structure for Vehicular Ad-Hoc Networks
[183]	Tracing the source of fake news using a scalable blockchain distributed network
[184]	Lightweight Blockchain Remote Mutual Authentication for AI IoT Sustainable Computing Systems
[185]	Lightweight Blockchain Security Protocol for Secure Storage and Communication in SDN-Enabled IoV
[186]	LBTM: A lightweight blockchain-based trust management system for social internet of things
[187]	LBlockchainE: A Lightweight Blockchain for Edge IoT-Enabled Maritime Transportation Systems
[188]	Identification of End-User Economical Relationship Graph Using Lightweight Blockchain BERT Model
[189]	Scalable Lightweight Blockchain-Based Authentication Mechanism for Secure VoIP Communication
[190]	LBSS: A Lightweight Blockchain-Based Security Scheme for IoT-Enabled Healthcare Environment
[191]	Research on Lightweight Blockchain Technology Based on Edge Computing
[192]	Lightweight Blockchain Framework For Medical Record Data Integrity
[193]	Improving IoT Data Security and Integrity Using Lightweight Blockchain Dynamic Table
[194]	Lightweight Blockchain-Based Scheme to Secure Wireless M2M Area Networks
[195]	A Case Study of Multi-Hop Clustering Algorithm Based on Spectral Classification Using Blockchain
[196]	Securing Internet of Things Environment using Lightweight Blockchain Approach
[197]	Lightweight Direct Acyclic Graph Blockchain for Enhancing Resource-Constrained IoT Environment
[198]	DAG blockchain-based lightweight authentication and authorization scheme for IoT devices
[199]	Agricultural lightweight embedded blockchain system: a case study in olive oil
[200]	An artificial intelligence lightweight blockchain security model for security and privacy in IIoT systems
[201]	An extended lightweight blockchain based collaborative healthcare system for fraud prevention
[202]	Lightweight Blockchain-empowered Secure and Efficient Federated Edge Learning
[203]	Lightweight Blockchain-Based Secure Spectrum Sharing in Space-Air-Ground Integrated IoT Network
[204]	Lightweight Blockchain-Based Architecture for 5G Enabled IoT
[205]	Review and Development of a Scalable Lightweight Blockchain Integrated Model for IoT Applications
[206]	Lightweight-BIoV: Blockchain Distributed Ledger Technology for Internet of Vehicles
[207]	Secure PBFT Consensus-Based Lightweight Blockchain for Healthcare Application
[208]	A Lightweight Blockchain Framework for secure transaction in resource constrained IoT devices
[209]	A Lightweight Blockchain and Fog-enabled Secure Remote Patient Monitoring System
[210]	TinyLedger: A Lightweight Blockchain Ledger Protocol for the MEC Network

One hundred thirty-six papers were excluded due to the following reasons:

- The paper is not focused on blockchain or other DLT architecture: 94 papers
- The paper does not include an evaluation section: 16 papers

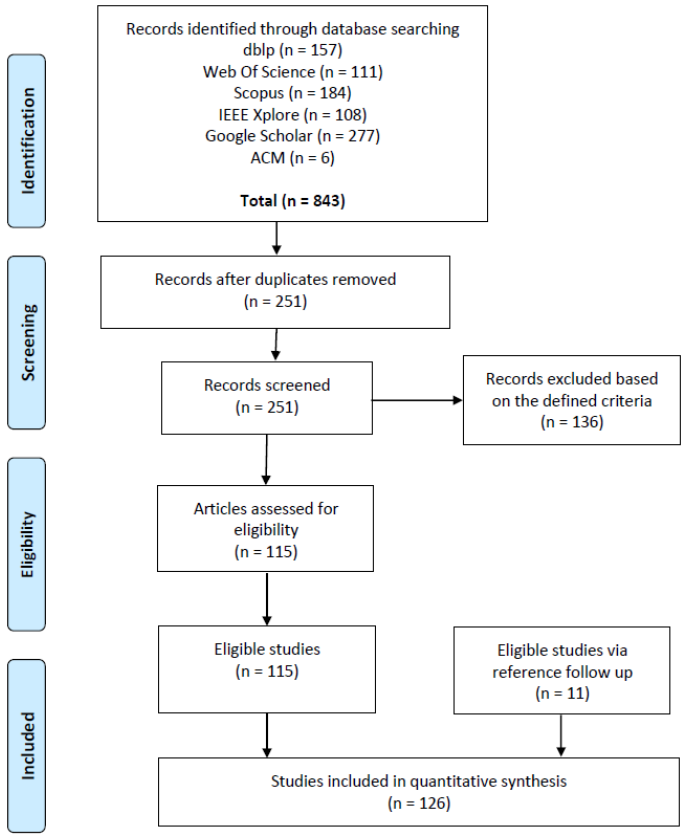

Fig. 3.2 PRISMA flow diagram

- It is not a research paper, or it is a review paper: 14 papers
- The paper is not written in English: 6 papers
- The paper is not available on the internet: 6 papers

Figure 3.3 illustrates the number of qualifying papers published annually since 2017, marking the publication date of the oldest study in the included research. As evident in the displayed graph, there has been a substantial rise in the number of papers focusing on designing DLT architectures for IoT environments over recent years. However, it is noteworthy to mention that in 2023, there seems to be a slight deceleration in the momentum of these publications. Several factors might account for this slowdown. The surging interest and developments in AI have understandably captivated much of the research community's attention, potentially diverting focus and resources away from DLTs. Additionally, some cryptocurrencies, which served as the base appli-

cations for DLT technologies, have recently encountered challenges and controversies. These events might have cast a shadow on the general perception and enthusiasm for DLTs, leading to a cautious approach in academic exploration and subsequently a dip in the number of publications.

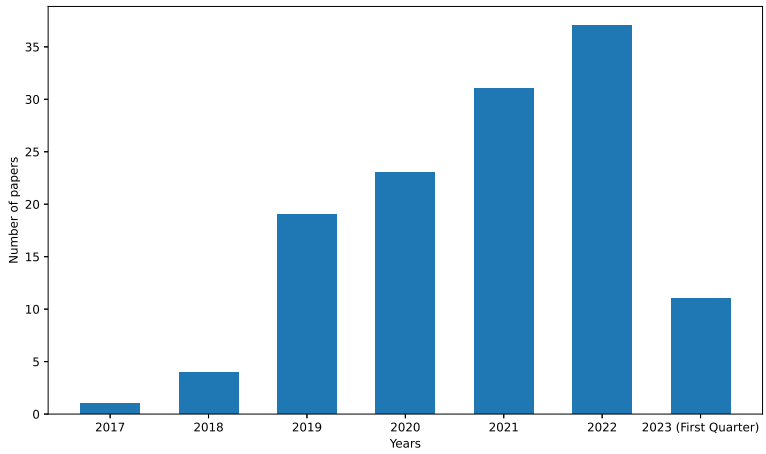

Fig. 3.3 Number of eligible papers published each year

The data extraction methodology from the included papers was defined following the research questions of this study and other possible relevant information. The extracted data are as follows:

- The author(s) name, the title, the publication year, the language the reference and the type of the paper.
- The field of application (e.g., industry, smart home, etc.).
- Main characteristics of the proposal: DLT type, structure, consensus and storage.
- Lightweight aspects: architecture, storage, consensus and cryptography.
- The evaluation process of each proposal: implementation method and evaluated metrics.
- Possible research opportunities for the future.

3.5 Results

In this section, the results of the collected data addressing the research questions that have been defined in Section 3.4 are presented.

3.5.1 RQ1: What are the most common fields of application for blockchain architectures?

The studied proposals have been designed for a wide variety of areas. However, the majority of the proposals are generic, which means that they can be applied in various (or any) domains.

Out of 126 studies reviewed, several standout papers were identified that represent key developments in lightweight DLT solutions tailored for specific application domains. In the context of the industrial field, or Industry 4.0, from 19 papers, five stood out as particularly influential [23] [99] [111] [126] [128]. These studies reflect meaningful contributions to the integration of DLTs in the industrial sector.

Noteworthy contributions were also identified in other areas. Papers by authors [88] [89] [90] provided pioneering insights into DLT architectures for smart homes. Additionally, research by [116] [127] [129] offered valuable perspectives on their application in smart vehicles. Other significant works include studies addressing the healthcare domain [98], the Internet of Drones (IoD) [91], the underwater IoT [94], and autonomous transaction settlement for e-commerce [95].

Notably, ten papers have focused on the crucial area of blockchain-based smart grids, whereas three works addressed the burgeoning field of 5G mobile networks.

While the remaining studies (n = 45) presented generic proposals, the aforementioned works stand out for their substantial contributions to their respective fields. As a result, the most commonly targeted fields where blockchain (and other DLT) solutions have been designed include:

- Generic domain (n = 45)
- Industry (n = 19)
- Smart vehicles (n = 15)
- Smart homes (n = 13)
- Smart grid (n = 10)
- Healthcare (n = 9)
- Other (n = 15)

3.5.2 RQ2: What characteristics do the proposals have?

In this subsection, a classification and comparison of the main characteristics of the reviewed proposals are presented. The characteristics that were gathered from the reviewed papers are as follows:

- The type of the blockchain (or other DLT type) in terms of access control.
- The structure of the blockchain (or other DLT type) architecture.
- The consensus protocol.
- The storage approach of the proposed architecture.

Below a summary of the characteristics that were reviewed for each of the included papers and an assessment of the gathered information is provided.

DLT type

There are two main types of blockchain (or DLTs) in terms of data access: permissioned and permissionless. Seventy-eight proposals were specifically designed as permissioned (n = 56) or permissionless (n = 22) blockchains, whereas five use both types in the same framework. In addition, 15 proposals were not designed for a specific type of blockchain; thus they could be used in both permissioned and permissionless environments (i.e., "any" type). Figure 3.4 shows the distribution of the characteristics.

Note that in many proposals, the type of the blockchain is not mentioned.

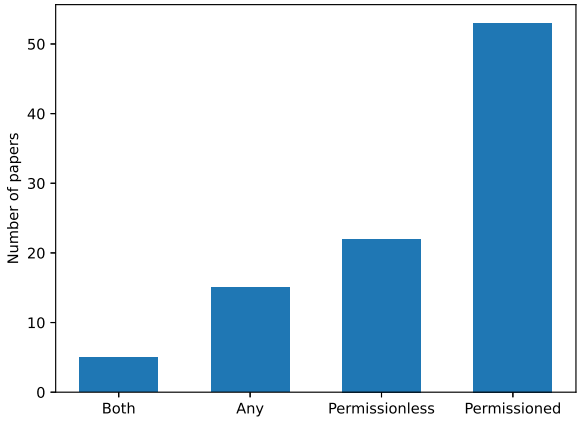

Fig. 3.4 Most used blockchain types

Structure

Originally, all the nodes in a DLT network could take the role of miners/validators while storing the entire chain. This type of structure can be defined as the "classic" structure. In resource-constrained environments, this type of structure is not usually possible [128]. Therefore, 49 authors divide the network into different layers of devices that have different capabilities and roles. In addition, the clustering method, where clusters of nodes are maintained by a cluster head, is also common, as it is used in 20 proposals. Both approaches (layering and clustering) can also be combined, as can be seen in 10 works. Another approach is the Directed Acyclic Graph (DAG), which is a structure used in a different type of DLT than blockchain and was firstly introduced by IOTA [211]. The DAG architecture was used in only seven works. Finally, 28 authors maintain the "classic" one-layered blockchain structure in their proposals. Figure 3.5 shows the distribution of the structure found in the reviewed papers.

Note that some proposals where the structure could not be determined due to the lack of information or the incompatibility of the type of proposal with this categorization.

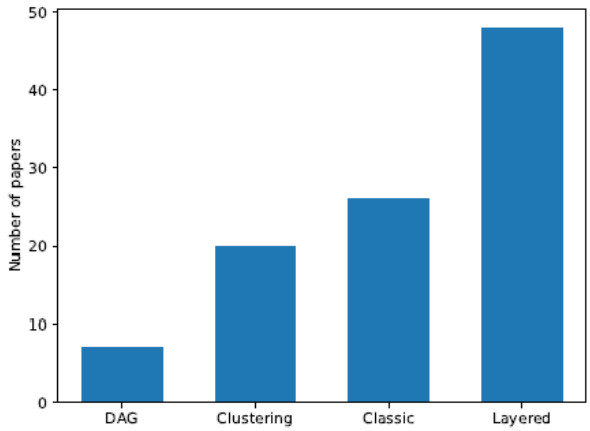

Fig. 3.5 Most used DLT structures

Consensus

Within the results of the analysis, the consensus algorithms can be divided in two groups:

1. **Custom-made consensus algorithms.** A custom consensus algorithm can be defined as an algorithm that was specifically developed for the proposed framework

and was not used in any other framework or system. In 25 papers, we can find different custom algorithms that are randomness, vote, time, trust or location-based.

2. **Generic consensus algorithms.** A generic consensus algorithm can be defined as an algorithm that was not specifically developed for a specific framework research study, or is applied in multiple frameworks or systems. In 43 papers, generic lightweight consensus algorithms such as Proof of Stake (PoS), Practical Byzantine Fault Tolerance (PBFT), Proof of Elapsed Time (PoET), Raft, Proof of Authority (PoA) and Proof of Capacity (PoC) are present.

The original consensus algorithm of blockchain is the PoW [212]. However, this algorithm is well known for its low efficiency and high resource requirements [213], which makes it unfeasible for resource-constrained devices. Therefore, most authors employed more efficient consensus algorithms when building lightweight DLT solutions. These algorithms are as follows: PoS, PBFT, PoET, PoA, PoC, Proof of Reputation (PoR), Raft and other custom-made consensus. All of the alternative consensus algorithms that were used in the reviewed papers were designed to overcome the drawbacks of the PoW algorithm in resource-constrained environments. Thirteen authors presented an enhanced version of the PoW algorithm rather than implementing a novel algorithm. Figure 3.6 shows the distribution of the discussed types of consensus in the reviewed papers.

Note that, in many proposals, the consensus algorithm is not mentioned.

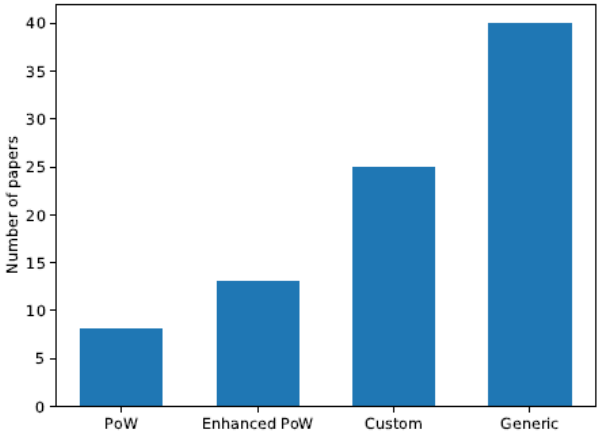

Fig. 3.6 Most used types of consensus protocols

Storage

Besides the consensus algorithm, storage is also a major issue in the DLT-based IoT environments [100]. This issue can be easily tackled in some fields where historical data are not important and therefore are stored temporarily. However, most of the time, this is not the case. Specifically, data storage can be addressed as follows:

- **On the blockchain (On-BC).** In 63 proposals, the data are kept inside the ledger. However, usually, lightweight IoT devices do not have sufficient storage space to keep the whole ledger. Therefore, 21 authors propose layered architectures where the data are stored in specifically designed storage nodes or layers within a DLT.

- **Cloud.** Sixteen authors combined Cloud Computing with blockchain in order to tackle the storage issue. In the framework that is presented in [89], the authors assume that a smart home user already has a Cloud account such as Dropbox. M. A. Uddin *et al.* [94] are the only authors that propose a cloud-based blockchain rather than just Cloud storage. They claim that this type of blockchain is the most optimal choice for the high processing and storage requirements of IoT.

- **Off-chain.** Ten authors proposed architectures where the data are stored off-chain (e.g., in a local server or database). In this approach, the only data that has to be stored on the shared ledger are its hashes in order to assure its integrity. However, storing data off-chain does not assure its availability.

Figure 3.7 shows the distribution of the type of data storage in the reviewed papers.

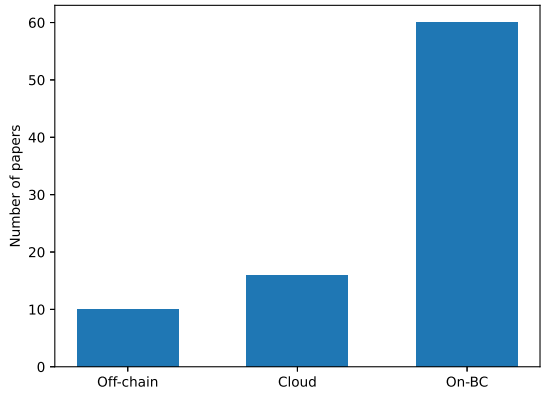

Fig. 3.7 How data is stored in lightweight blockchain

Note that, there are three proposals where the storage approach could not be determined due to the lack of information or the incompatibility of the type of proposal with the categorization of this characteristic.

3.5.3 RQ3: Which aspects do the authors optimize in their architecture?

This subsection studies the lightweight aspects of the reviewed proposals.

The studied aspects are as follows: consensus, storage approach, architectural structure and cryptography.

An evaluation criterion for each one of the considered aspects was established: consensus, storage, architecture and cryptography, as shown below.

Consensus

In permissioned networks, resource-intensive consensus such as PoW is not necessary [214]. Thus, this work considers that a proposal is lightweight in terms of consensus if it is a permissioned framework that uses:

- A custom vote, time, trust or location based algorithm.
- A generic consensus algorithm that was designed as an efficient alternative to PoW such as: PoS, PBFT, PoET, Raft, PoA, PoC and PoR.

Furthermore, an enhanced version of the PoW algorithm could also be considered as lightweight if the authors provide enough evidence on its suitability for resource-constrained devices.

Storage

As mentioned in Section 3.5.2, one of the main features of blockchain (and other DLTs) is the fact that the ledger is replicated in all devices involved in the network. Thus, if attackers want to forge the data, they must hack the majority of devices [100]. However, a resource-constrained device cannot maintain the ledger continuously because of its low capabilities. Therefore, a solution can be considered to be lightweight in terms of storage if:

- The data are stored temporarily on the ledger.
- The data are stored outside the ledger (e.g., on the Cloud or in an external database or server).

- There is enough evidence that the size of the data or blocks is reduced so that the storage of the ledger is feasible on resource-constrained devices.
- The data are only stored in a specific layer or storage nodes within the ledger.

Architecture

Resource-constrained devices are unable to participate and maintain a blockchain or other DLT network [23]. Therefore, an efficient architecture must divide the network in various layers and/or clusters that give the involved devices different tasks according to their capabilities.

Cryptography

Blockchain and most DLTs are strongly based on cryptography [215]. However, cryptography processing in resource-constrained devices is not straightforward. Therefore, a solution can be considered as cryptographically lightweight if there is strong evidence of a significant performance improvement related to the cryptographic part of DLTs for resource-constrained environments.

As it can be seen in Figure 3.8, based on the defined criteria, 74 proposals optimized the consensus, 63 designed a lightweight architecture, 49 a lightweight storage and 11 a lightweight cryptography. Only four proposals are optimized in all aspects: consensus, storage, architecture and cryptography. The rest of the work could not be evaluated in this regard.

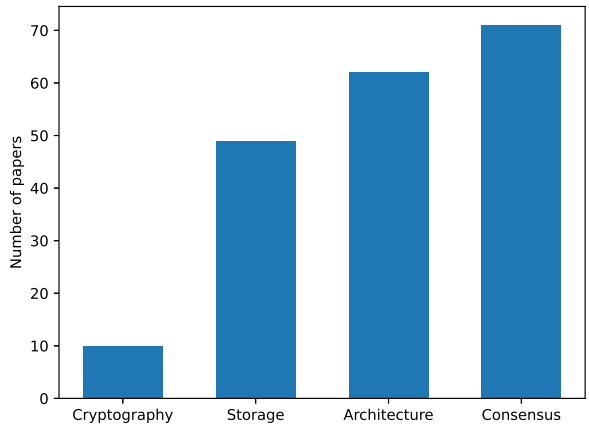

Fig. 3.8 Lightweight aspects of the proposals

3.5.4 RQ4: How are the proposals evaluated?

This subsection studies the evaluation of the reviewed papers. It reviews the method of implementation of each paper and the evaluated metrics.

Implementation method

Twenty-four authors used a high level programming language for the implementation such as Python (n = 13), Java (n = 6), C/C++ (n = 3), JavaScript (n = 1) and iOS Swift (n = 1). Thus, Python, Java and C/C++ are the most commonly used programming languages for DLT development. Twenty-three authors implemented their proposals in specific DLT development platforms such as Hyperledger (n = 11), Ethereum (n = 10) or Multichain (n=2). Platforms such as Hyperledger Ethereum or Multichain offer great possibilities for implementing DLTs as they are suited for permissioned networks that include lightweight consensus. Twenty-two authors used generic simulators such as the NS-3 network simulator (n = 10), Cooja (n = 4), Matlab (n = 4), Colored Petri Net (n=1) or custom made simulators such as "ZeroCaloSimu" (n=1) or "BlockLite" (n = 2). Finally, 29 authors did not provide information on how their solution was developed. Hence, in that case, the implementation parameter was marked with a "not available" abbreviation (N/A). Figure 3.9 shows the distribution of the implementation methods that were used in the reviewed papers.

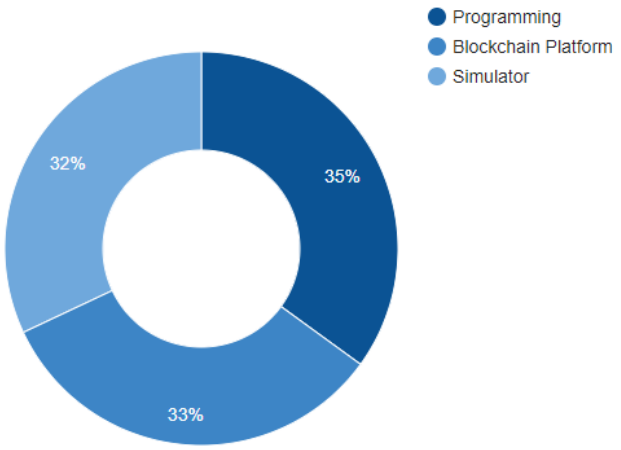

Fig. 3.9 Implementation methods

Evaluated metrics

This sub-subsection studies the metrics that were evaluated in the reviewed papers. This review is mostly focused on the performance of DLTs; hence, security evaluations are omitted. As it can be seen in Figure 3.10, the authors of the reviewed papers have evaluated a wide range of performance metrics. Each author evaluated different metrics based on different criteria. The authors mostly focus on evaluating metrics that are related to their proposal's strengths and aimed improvements. The gathered evaluated metrics can be framed in the following categories:

- **Computational. (n = 82)** The metrics that are related to the computational resources such as the CPU, the memory, etc.
- **Blockchain (or otherDLT). (n = 78)** The metrics that are related to the DLT transactions, blocks and consensus.
- **Network. (n = 69)** The metrics that are related to the network communication, such as bandwidth, latency, etc.
- **Storage. (n = 53)** The metrics that are related to the data storage.
- **Energy. (n = 33)** The metrics that are related to the energy or power consumption.
- **Cryptography. (n = 11)** The metrics that are related to the cryptography.

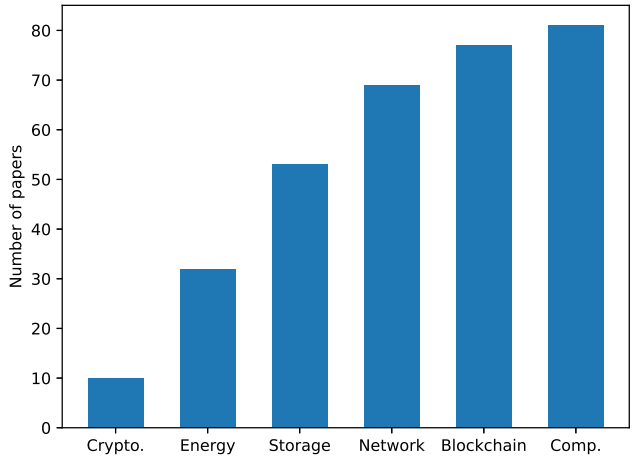

Fig. 3.10 Evaluated metrics

3.6 Discussion

This section analyses the results of the study and give some insights. It also discusses several research opportunities that have been identified during the review process. This section is divided in three subsections:

- Section 3.6.1 is related to the fields of application of lightweight DLTs. Specifically, it addresses RQ 1.
- Section 3.6.2 is related to the technical aspects of lightweight DLTs. Specifically, it addresses RQs 2 and 3.
- Section 3.6.3 is related to the evaluation of lightweight DLTs. Specifically, it addresses RQ 4.

3.6.1 Fields of application - Industry 4.0

In the presented analysis, it is evident that many authors promote DLTs, especially blockchain, as the panacea to the challenges encountered across various domains. However, one cannot overlook the fact that implementing DLTs in resource-constrained environments is not always straightforward. Such implementation comes with its unique set of obstacles that necessitate extensive research. Nevertheless, this examination indicates that it is possible to effectuate significant enhancements to DLT applications in numerous sectors, reinforcing its position as a technology with tremendous potential. Specific challenges, such as security concerns and issues of centralization, seem to find their most promising resolutions in DLTs. Intriguingly, while many of the solutions that have reviewed possess a generic nature (meaning they can be adapted across multiple domains), it is evident that fields based on the Internet of Things (IoT) grapple with ubiquitous challenges like scalability and potential single points of failure due to centralized structures.

Yet, it is crucial to highlight that while IoT-related sectors share some challenges, many application fields come with their bespoke set of issues. This section pivots the focus primarily towards the outcomes related to Industry 4.0, the primary domain of interest of this thesis. The intersection of DLTs with Industry 4.0 presents a unique fusion of challenges and solutions, and the intent is to delve deeply into this confluence.

Industry 4.0 - IIoT

The intersection of blockchain and DLTs with IIoT presents a complex challenge, requiring comprehensive solutions that address inefficiencies in DLTs and consider the

broader industrial scenario. Focusing solely on solving individual DLT-related issues may lead to sub-optimal architectures for Industry 4.0. A broader view should be taken when designing solutions to holistically address IIoT needs in the industrial environment.

Efforts to distribute the workload of DLTs across layers, such as those found in Edge Cloud or resource-layered architectures, offer potential benefits for IIoT. However, these approaches may not be entirely effective in the industrial context. The complications arising from distributing the workload should be carefully considered and addressed in future research to improve the applicability of layered DLT architectures.

There is potential in integrating other DLTs, such as DAGs, to complement or enhance the use of blockchain for IIoT. It is crucial not to overlook these alternatives when designing more efficient architectures for Industry 4.0. Embracing a broader range of DLTs could help provide more flexible and efficient solutions for IIoT.

The inclusion of energy-consuming processes like mining remains a significant limitation in some of the discussed approaches. This issue should be addressed in future research to promote more sustainable and efficient solutions for IIoT in Industry 4.0. Creating energy-efficient architectures is essential for the long-term viability of DLTs within industrial environments.

In summary, while progress has been made in adapting DLTs to IIoT, a more holistic approach is needed to address the unique requirements of IIoT in Industry 4.0 efficiently. Future research should focus on building architectures that seamlessly integrate DLT into the industrial landscape, enabling the potential benefits of these technologies to be fully realized.

DLT interoperability and oracles

There is a notable trade-off between achieving interoperability and maintaining performance efficiency, scalability, and security. Many of the proposed solutions incur high-performance costs, low scalability, or network latency, which may limit their practical application.

While the use of notaries, oracle services, or relay architectures offers potential solutions for interoperability, these approaches can introduce new vulnerabilities or single points of failure into the network. These centralized elements are at odds with the fundamental principles of decentralization that underpin DLTs.

There appears to be a need for greater consideration of the specific requirements of industrial environments, such as the heterogeneity of data and the need for secure methods of introducing external data into smart contracts.

Many of the proposed solutions focus on achieving interoperability between dif-

ferent blockchains, while other promising DLTs, such as DAGs, have not been fully explored. Future research should take a more inclusive approach to encompassing a variety of DLTs.

While the proposed approaches offer innovations, they also have clear limitations that need to be addressed. Future work should focus on developing more comprehensive, efficient, and secure interoperability solutions that can accommodate the unique requirements and challenges of Industry 4.0 ecosystems.

Modern industry monitoring

It is clear that modern industrial monitoring systems under the Industry 4.0 framework involve the integration of various technologies and approaches. These schemes cover a range of functionalities, such as real-time data acquisition, automation, and comprehensive monitoring of industrial equipment.

Several important trends emerge from these monitoring schemes. First, automation is a key factor in the presented systems. By removing manual interference and continuously registering signals, the schemes significantly improve efficiency within industrial settings.

Second, the use of wireless technologies and embedded hardware for remote monitoring is a prevalent feature among the schemes. The ability to monitor critical parameters like energy consumption, temperature, and CO2 levels from a distance is invaluable in enhancing safety and efficiency.

Third, the integration of IoT platforms for real-time visualization is a key aspect of modern monitoring systems. By providing a platform for remote visualization of data, these systems can contribute to more informed and timely decision-making processes in industrial environments.

Fourth, dealing with the challenges of large data volumes and integrating key technologies such as Wireless Sensor Network (WSN) and Radio Frequency Identification (RFID) into monitoring systems is crucial. These technologies are essential for robust data processing and integration capabilities, which are vital for handling the complexities of modern manufacturing environments.

Fifth, the use of IoT analytics platforms for data analysis and generating reports is a common trend. The ability to identify anomalies, offer improvement suggestions, and even incorporate machine learning for enhanced prediction and mitigation of failures shows the growing importance of advanced analytics in industrial monitoring.

Finally, the emergence of augmented reality and mobile application tools for monitoring systems is an innovative development. Such systems offer precise monitoring, maintenance scheduling, and increased interoperability and communication, providing

valuable data for further analysis.

In conclusion, the monitoring schemes in the context of Industry 4.0 demonstrate a trend toward automation, wireless technologies, real-time data visualization, and advanced analytics. These features enhance efficiency, safety, and informed decision-making within modern industrial environments.

Industrial DLT design frameworks

The exploration of DLT technologies within supply chain networks and organizational systems has led to the development of several valuable frameworks and methodologies. These tools provide essential criteria for evaluating and selecting suitable DLT platforms for specific use cases.

One such framework offers an extensive set of criteria across multiple dimensions, providing comprehensive guidance for assessing DLT platforms, particularly in supply chain networks. This framework's thorough approach makes it a useful reference for organizations looking to implement DLT technology in various sectors.

An automated decision-making framework provides a systematic benchmarking process, aiding in the selection of alternative platforms based on specific requirements and preferences. By matching necessary requirements with established quality features, this tool facilitates informed decision-making about suitable DLT platforms.

Another comprehensive framework highlights the importance of a hierarchical approach to criteria selection and weighting. It emphasizes crucial criteria like cost, speed, privacy, functionality, and developer availability, offering detailed guidance for evaluating potential DLT platforms.

A selection methodology highlights the significance of aligning the DLT platform's attributes with the specific needs of an enterprise system. By emphasizing the difference between technology-based selection methodologies and domain-specific processes, this approach provides valuable insights for organizations looking to select the right DLT platform.

An overview of DLTs applications within Industry 4.0 showcases its potential to enhance various aspects of supply chain management, data security, M2M communication, and cybersecurity. This broad exploration serves as a valuable reference for researchers and developers interested in creating DLT frameworks for Industry 4.0.

In conclusion, the tools and insights from the research contribute to the assessment and selection of DLT platforms within supply chain networks and other organizational systems. These comprehensive frameworks, criteria, and methodologies facilitate informed decision-making for developers and organizations interested in implementing DLTs.

Industry 4.0 smart contracts for business applications

The discussion of the various studies demonstrates that while smart contracts offer transformative potential in various industrial contexts, there are several challenges that need to be addressed. One recurring theme is the need for a more comprehensive set of data accessible to smart contracts. Many of the studies narrowed their scope to a single business case, and their smart contracts were restricted in accessing external information. This highlights the importance of creating smart contract systems capable of accessing a wide range of data sources to enhance their flexibility and applicability across different business cases.

Another critical issue is the need for enhanced security in smart contracts. Authors utilized AI trained models for anomaly detection in a decentralized traffic control platform. While this approach offers valuable insights into achieving better security in smart contracts, it relies heavily on AI, which might not be a feasible or necessary solution for all scenarios. This suggests that there is a need for smart contract platforms focusing on inherent security mechanisms without necessarily relying on AI.

The discussion also emphasizes the importance of scalability and performance in the development of smart contract platforms. Many works proposed a smart contract middleware for achieving secure, decentralized industrial communication, but the block times were found to be too high, rendering the approach impractical. This indicates the critical need for developing smart contract systems that are both scalable and high-performing.

Moreover, there is a recurring theme of utilizing smart contracts beyond mere access control. Many authors designed smart contract-based access control mechanisms for Industry 4.0, where data was stored in a decentralized database with secure smart contract-based access. While this approach is valuable, it highlights the potential for expanding the usage of smart contracts to automate more intricate business processes.

Finally, several studies focused on specific use cases, indicating the need for a more versatile, use case-agnostic smart contract platform. Developing such a platform would provide greater flexibility and applicability across a broader range of scenarios.

In conclusion, the insights from the presented studies underscore the importance of designing smart contract platforms that are versatile, scalable, and can access diverse data sources. There is a clear need for developing smart contract systems that focus on inherent security and decentralization mechanisms. Furthermore, there is significant potential for utilizing smart contracts beyond mere access control towards the automation of intricate business processes. Creating flexible and robust smart contract platforms for broad applicability across various business cases and scenarios is essential to unlocking the full potential of smart contracts in industrial contexts.

3.6.2 DLT Architectures Technical Aspects

This work identified very few solutions that are lightweight in all of the studied aspects: consensus, architecture, storage and cryptography. Thus, there is a clear need to design complete lightweight DLT frameworks in order to fulfill the needs of resource constrained environments.

When it comes to lightweight blockchain and other DLTs, the majority of researchers think about the computational burden of blockchain in the first place. Blockchain offers major security and privacy features to networks that are composed of untrusted devices. However, these advantages come at a huge cost in terms of computational burden.

According to most of the authors, the part of blockchain that mostly causes its computational burden is the consensus algorithm. In consequence, many alternatives to the original PoW consensus algorithm of blockchain have been proposed. According to the results of the study, vote-based consensus algorithms are highly efficient and secure, whilst the PoW algorithm is the least efficient. It is worth mentioning that improving the PoW algorithm is also a studied option. However, enhancing the performance of the consensus algorithm could have a serious impact on the security of blockchain. That is why most authors design permissioned blockchain architectures for IoT. In a trusted environment, the security features of the consensus algorithm can be reduced in order to lower its computational burden.

Another effective method of reducing the computational burden of blockchain is to design layered and/or clustered architectures. Dividing an architecture into various layers or clusters prevents resource-constrained devices from performing heavy computational tasks such as mining. However, this approach also has a negative impact on some benefits of blockchain. Ideally, all devices should participate in the blockchain network in order to assure maximum security and trust.

Thus, the main conclusion on the computational burden issue is that further research is required. There is a considerable need to develop more lightweight consensus algorithms without sacrificing security. Also currently, designing layered architectures where IoT devices do not have to perform heavy tasks is an optimal approach.

The second concern of the researchers that work on lightweight DLT solutions is the network overhead. In blockchain and many other DLTs, all the transactions that occur in the network must be replicated in all nodes. In addition, lightweight consensus that is based on voting also carries an enormous communication burden. For example, the PBFT consensus needs to constantly exchange information regarding blocks validation between all the nodes of the network. That is why the performance of PBFT dramatically decreases when the number of nodes is high (i.e., more than 20) [214]. One of the most effective ways to reduce the network burden in blockchain is presented in [23].

The authors observed that during blocks verification, the information broadcast by peer nodes overlapped. Hence, they designed a lighter block structure named LightBlock. This approach reduces the necessity of sending the entire data to the other nodes more than one time. This approach reduced the network overload by over 90%. However, reducing the network burden in distributed systems while maintaining the full availability and integrity of the data is still a major issue that needs further research.

One of the greatest drawbacks of blockchain and other DLTs is the low throughput. Bitcoin can only process seven transactions per second [212], whereas conventional payment systems like VISA or PayPal can process thousands. The low throughput of blockchain is not only a major issue in financial applications. IoT generates thousands of exabytes annually [216], and all that data has to be processed rapidly. The throughput is another aspect of blockchain that is strictly tied to the consensus algorithm. The heavy consensus process of blockchains greatly reduces their throughput. The most remarkable mechanism that has been proposed in order to improve the throughput of blockchain is the reputation-based consensus. One of the most effective reputation consensus is proposed in [89]. In this type of consensus, the nodes that have a good reputation are able to generate transactions at a much faster rate. This is because when trust is created, the verification process decreases for the nodes that have proved to be trustworthy. However, one of the major drawbacks of reputation-based consensus is that a trusted (i.e., permissioned) environment is required. Therefore, improving throughput in permissionless DLTs is still a major issue that needs further research.

One of the main features of blockchain and many other DLTs is the fact that the ledger is replicated in all devices involved in the network. Thus, if attackers want to forge the data, they must hack most of the devices in the network. However, a resource-constrained device cannot maintain the blockchain continuously because of its low capabilities. Specifically, due to insufficient storage capacity, these types of devices cannot assure DLT property of immutability [100]. According to the results of the study, there are three main approaches for lightweight DLTs storage:

- Storing the data in the ledger, but not on all devices. This approach is very typical in layered architectures, where the data are stored in nodes that have sufficient storage. However, this approach separates the lightweight devices from the DLT network itself. As it was mentioned before, ideally, all devices should fully participate in the DLT network.

- Storing the data off-chain is a simple yet effective method of reducing the storage burden in DLT. In this approach, the only data that has to be stored in the DLT are its hashes in order to assure its integrity. However, storing data off-chain does

not assure its availability, which is a major issue. Therefore, it is recommended to use this approach in environments where data loss is not a major concern.

- Cloud computing is another effective method of reducing the storage burden in DLTs. This method is very similar to the previous one. However, cloud storage is maintained by a third party. Thus, this approach is recommended to be used only if the privacy and the availability of the data are not critical.

In conclusion, there is a clear need to further research the integration of Cloud computing with DLTs in order to deliver safe, lightweight storage for resource-constrained environments. Furthermore, assuring the availability of the data in an off-chain storage approach is also a great challenge that requires further research. Nonetheless, novel approaches that would reduce the storage burden of on-chain data would be the most appropriate method of improving this aspect.

Energy consumption is the least aspect that authors mention when working on lightweight DLTs. However, this aspect has a huge impact on our world. Thus, it is not less important. According to [217], Bitcoin mining consumes the same amount of energy as the entire country of Denmark. Nevertheless, the huge energy consumption of blockchain and other DLTs not only involves environmental issues. Millions of IoT devices run on batteries [218], making most DLTs unfeasible for a great part of lightweight devices. Energy consumption is mostly tied to the consensus algorithm. Therefore, improving the consensus algorithm also has a positive impact on energy consumption. For example, the authors in [23] propose a "green" consensus algorithm that reduces mining, with the specific purpose of reducing the energy consumption of blockchain in industrial environments. Many authors completely removed the mining process of the consensus in order to reduce energy consumption. However, as mentioned before, removing mining could drastically reduce the security of the blockchain. This is why the most efficient consensus algorithms are available only in permissioned networks. Therefore, further research on efficient consensus for permissionless blockchain and other DLTs is recommended.

Very few authors focused on cryptographic improvements. Cryptography is a core feature of DLTs, especially blockchain [219]. However, cryptography incurs a major burden, especially in lightweight IoT devices. Therefore, some of the reviewed papers aimed at reducing the burden that cryptography causes in IoT. In [126], the authors address the performance and energy consumption of the hash function in the mining process. They propose a novel mechanism that can change the hash algorithm used for mining by adjusting to the network traffic. The work [220] addresses a similar problematic regarding the performance of cryptographic functions in DLTs. The work in [220] analyzes the implementation of Federated Learning (FL) algorithms in blockchain for

IoT schemes. FL algorithms improve blockchain-based IoT architectures by adding privacy and by further reducing overhead. Similarly, the authors in work [125] improve the used algorithms from the proposal presented in [89] achieving better security and performance results. However, the cryptography enhancement has received too little attention from the researchers and that there is room for more improvements. Novel lightweight cryptographic functions for DLTs need to be developed. Furthermore, it is also important to take into account quantum computing, which can pose a major threat to the security of DLTs [221].

Another approach for lightweight DLTs that is worth mentioning is the DAG structure. IOTA introduced this type of DLT aiming at IoT environments. However, this framework is not completely decentralised yet, since it has a centralised coordinator. The coordinator is run by the IOTA Foundation in order to assure the security of the network. Currently, DAG based DLTs can be completely decentralised and secure only when there is a high volume of transactions. One highly relevant lightweight architecture based on DAG structure is presented in [129]. In this paper, the authors try to tackle the storage issue of DLTs by proposing a DAG network for vehicular social networks. In the proposed architecture, only recent data that is useful for the drivers is maintained in the ledger. Furthermore, the main ledger is divided into various topic groups, which also greatly reduces the storage requirements. One particular DAG approach is presented in [145], where the authors design a DAG architecture that is very similar to blockchain, thus maintaining its greatest drawbacks such as huge energy consumption due to PoW mining. However, this particular DAG structure offers much more throughput capacity than regular blockchains. In conclusion, DAG is a promising solution. However, this technology still has some important limitations and challenges, such as centralisation and security issues [211]. Furthermore, DAGs still require real-world validation in several IoT areas. Apart from DAG DLTs, there are several efficient blockchain solutions that are suitable for IoT; the Hyperledger ecosystem, with Fabric and Sawtooth as the most used blockchains, and other platforms such as R3 Corda or Ethereum 2.0 with the novel PoS scheme that was recently released. Hashgraph is also an emerging solution that offers great efficiency. However, this technology has not yet been consolidated.

Finally, according to the previous discussion, we can conclude that the most justifiable aspects that make a DLT "lightweight" are as follows: efficient consensus algorithm, external storage and efficient cryptographic implementations. Consequently, the aforementioned characteristics guarantee low energy consumption, low network overhead, low computational and storage burdens, and overall high throughput capacity.

3.6.3 DLT Architectures Evaluation

The evaluation of the reviewed papers has been analysed. Specifically, the implementation methods and the evaluated metrics of the proposals were analysed.

Implementation methods

This analysis shows that there is a wide range of implementation methods for deploying DLT networks. There is a clear lack of a simple, universal and standardised testing and evaluation platform for DLTs. Furthermore, the conducted experiments could not accurately reproduce the system behavior in a real-world environment due to the following reasons:

- Developing a DLT framework proof of concept from scratch using a high-level programming language is not a simple task, and there is no guarantee that it will provide reliable results.
- Available test environment might use different mechanisms from the real world implementation.
- Normally, only a small number of IoT devices are used.
- Simulations might not provide accurate results for all case scenarios.

Evaluated metrics

Each author focused on different metrics in order to validate their proposal. There are two main reasons for this; First, the authors focus on different problematic aspects of DLTs. For example, in [145] the authors claim that improving throughput makes this technology sufficiently suitable for IIoT, and therefore, only measure the transactions per second of their solution. On the other hand, the authors in [23] take more aspects into account and therefore include more metrics in their evaluation. Second, the fact that authors use many distinct platforms to perform their experiments also impacts the measured metrics. For example, Hyperledger comes by default with several tools that can be used for performance evaluation purposes, such as Hyperledger Caliper, whereas other platforms such as Ethereum only include metrics related to the blocks and transactions. Moreover, the existence of multiple evaluation environments developed from scratch provides infinite possibilities when defining evaluation metrics.

The results of this study have proven that lightweight DLTs must possess several key characteristics in order to be applied to IoT: low computational burden, low network overhead, low storage requirements, high throughput and high energy efficiency.

Therefore, a standardised metrics scheme for evaluating DLT solutions should be developed. Also, there is not clear what performance values are acceptable for a DLT to be considered "lightweight". For example, how many transactions per second are enough or can be acceptable for a DLT architecture for IoT? Establishing a consensus in this regard is an important challenge that needs to be addressed if standardised methodologies for lightweight DLTs are to be developed.

In conclusion, the high variability of the evaluated metrics in lightweight DLTs shows that it is necessary to develop a systematic and standard methodology in order to evaluate lightweight solutions. This would accelerate and facilitate the development and adoption of DLTs in various fields.

3.7 Summary and Conclusion

This chapter presents the state-of-the-art analysis that has been conducted thorough the exploration and analysis of 126 DLT architectures, all published since 2017, spanning a wide array of sectors. This encompassed categorizing the unique features of these DLTs, their lightweight components, and their respective evaluations. It also delved into the identification of existing weaknesses and suggested avenues for future research.

The conclusions drawn from this review indicate an upward trend in the popularity of DLT-based solutions. Each year has witnessed a significant surge in this interest, with the research papers examined covering a remarkable range of applications. They particularly focused on addressing unique challenges associated with implementing DLTs in resource-constrained environments, such as those found within IoT and Industry 4.0 settings in general.

Despite the substantial number of proposed solutions in this domain, this study strongly emphasizes the need for further research. The balance between security and efficiency in DLTs is delicate; any compromise on security could compromise the inherent advantages of DLTs over other alternatives. This underscores the necessity for additional research aimed at enhancing DLT-based architectures, particularly with respect to the creation of interoperable solutions that can seamlessly function across different platforms.

This brings into focus the importance of DLT interoperability and the deployment of smart contracts, which hold the potential to automate and streamline processes across a multitude of sectors. In particular, traceability, one of the inherent strengths of DLTs, can be leveraged to improve data integrity and transparency in various use cases, including supply chains and data management in Industry 4.0.

As this thesis pivots towards the design of a comprehensive DLT architecture for

Industry 4.0, a scope that encompasses not only the IoT segment but all higher levels, the findings from this review offer a valuable foundation. The primary goal with this research is to ensure the security, immutability, and traceability of data, right from its generation through to business-level processing. This study findings suggest that DAG DLTs represent a promising, yet relatively unexplored, terrain in the DLT landscape. Characterized by lower energy consumption, zero fees, and high throughput, this structure offers a compelling alternative. Yet, the relevance of "classic" DLTs is not diminished. Vote-based or round-robin consensus algorithms, coupled with layered Edge architectures, are equally efficient and have broad applicability across many IoT contexts. Additional promising solutions revolve around easing the storage burden of the DLT, particularly through decentralized databases such as Interplanetary File System (IPFS) for actual data storage. Given the potential threat quantum computing poses to current DLT architectures, research in the area of post-quantum cryptography also shows promise.

In consideration of these conclusions, a formally positive outlook on the future application of DLT technologies in IoT oriented fields can be maintained, including Industry 4.0. This perspective is particularly attributed to the potential these technologies exhibit in augmenting security, ensuring data immutability, enhancing traceability, etc. The upcoming design of a comprehensive DLT architecture for Industry 4.0, as presented in this thesis, will build on the insights and opportunities identified in this systematic literature review. With an emphasis on interoperability, smart contract functionality, and data traceability, this work aims to leverage the strengths of multiple DLT architectures to foster a secure, efficient, and transparent digital infrastructure for Industry 4.0.

While this state-of-the-art analysis has provided valuable insights into the status quo of DLT architectures and their applications in Industry 4.0, it is important to acknowledge its limitations. Given the expansive nature of DLT applications and the constant developments in this field, the study may not have covered all existing DLT architectures or explored every specific theme within the Industry 4.0 framework. For instance, areas like data homogenization, the use of smart contracts for business logic automation, and direct comparisons between different DLTs have not been extensively addressed.

Furthermore, there is still much room for research into specialized topics such as the integration of post-quantum cryptography, consensus algorithm optimization, and the creation of standards for interoperability between disparate DLT systems. This systematic review should, therefore, be considered a foundational piece upon which future research can build.

Chapter 4

A DLT-based Architecture for Industry 4.0

4.1 Overview

In this chapter we describe the DLT architecture proposal that is presented in this thesis as its main contribution. We deploy it on top of the Industry 4.0 case scenario that was detailed in Chapter 2. In the aforementioned scenario, several machine level objects (i.e., IIoT) form a production line, while several production lines form a plant, and several plants form a consortium. Thus, we intend to set up a DLT network at each level where a set of objects are located, in order to cover the whole industrial process from when the data is generated at the machine level up until the data are processed at the business level. Therefore, the proposed DLT architecture consists of three layers:

1. Data source DLT layer: at the production line level (several machines).

2. Bridge DLT layer: at the plant level (several production lines).

3. Business DLT layer: at the consortium level (several plants).

Section 4.2 provides an introduction to the architecture. Section 4.3 describes the first layer of the architecture, the Data Source Layer. Section 4.4 describes the second layer of the architecture, the Bridge Layer. Section 4.5 introduces a decision tree for designing business based blockchains in the frame of Industry 4.0. Section 4.6 describes the third layer of the proposed architecture, the Business Layer. Finally, Section 4.7 presents the validation of the architecture and Section 4.8 shows the summary and conclusions of the chapter.

4.2 Introduction

A multi-layer DLT architecture, designed to address the unique requirements and intricacies of data management and security at different levels within the Industry 4.0 case scenario is presented in Chapter 2. DLT technologies provide a secure, decentralized, and tamper-proof method for storing and sharing data, making it an ideal solution for addressing the data management challenges associated with modern industrial processes.

The proposed multi-layer DLT architecture aims to facilitate secure and efficient data handling across the entire industrial ecosystem, from the machine level to the consortium level of the presented industrial scenario. By integrating DLT networks at various stages of the industrial process, the architecture enables seamless communication and information flow, creating a connected, agile, and secure manufacturing environment. This holistic approach fosters trust and collaboration among stakeholders, promotes the integration and interoperability of industrial processes, and paves the way for the implementation of advanced Industry 4.0 solutions, such as smart contracts and decentralized applications.

The layers of the architecture are as follows:

1. **Layer 1 - Data source DLT layer:** This layer is situated at the production line level, which comprises several machines. Its primary role is to securely capture, store, and manage data generated by IIoT devices in real-time. This ensures data integrity and provides a tamper-proof record of machine-level activities, which is essential for traceability and auditability. The data source DLT layer enables real-time monitoring and control of individual machines and their interconnections. This facilitates efficient production line management and proactive identification of potential issues, such as machine failures or bottlenecks.

2. **Layer 2 - Bridge DLT layer:** Established at the plant level, this layer encompasses multiple production lines. Its primary function is to aggregate data from the first layer, homogenize and monitor it in order to make it exploitable and securely transmit it to the business DLT layer where it would be exploited for aggregated value and decentralized business agreements. Thus, this layer enables seamless communication and data exchange between different production lines within a plant, promoting optimal resource allocation and coordination among various production processes. The bridge DLT layer supports the integration of multiple plants and production lines, fostering a more connected and efficient manufacturing ecosystem. By leveraging the bridge DLT layer, plants can share

data securely and transparently, enabling cross-plant collaboration, benchmarking, and best practice sharing. This layer also facilitates the implementation of plant-wide performance monitoring and analytics, which can further enhance operational efficiency and competitiveness.

3. **Layer 3 - Business DLT layer:** Operating at the consortium level, which includes several plants and other business partners such as providers, clients, headquarters, etc. This layer's primary objective is to process, analyze, and manage data from multiple plants to support strategic business decisions and derive valuable insights. By leveraging the power of DLT, the business layer ensures data security, privacy, and traceability, allowing consortium members and external stakeholders to establish trust in the system. This layer enables the consortium to monitor and analyze performance metrics, identify trends, and uncover opportunities for improvement across the entire organization. Furthermore, the business DLT layer facilitates the implementation of smart contracts for automating business processes, such as supply chain management, product tracking, quality control, and regulatory compliance, enabling a wide range of business opportunities. This automation not only improves overall operational efficiency but also reduces costs, enhances transparency, and strengthens the consortium's competitive advantage.

In the subsequent sections, a comprehensive description of the three layers comprising the multi-layer DLT architecture is provided: the data source DLT layer, the bridge DLT layer, and the business DLT layer. A discussion on their primary functions, features, benefits, and potential use cases to demonstrate how they contribute to the realization of the Industry 4.0 vision and address the complex data management challenges facing today's industrial organizations is presented.

4.3 Layer 1: Data source layer

4.3.1 Introduction

The IIoT serves as a cornerstone of Industry 4.0, revolutionizing traditional industrial operations through the integration of smart sensors, robotics, Machine-to-Machine (M2M) communication, big data, and AI. Spanning various sectors, including manufacturing, energy, transportation, agriculture, and retail, IIoT optimizes production processes by enhancing customer experiences, reducing costs, and boosting efficiency. However, IIoT faces significant challenges, such as security, data privacy, and centralization, which hinder its full potential.

The emergence of the blockchain has presented a promising solution to IIoT challenges. Blockchain's inherent features, such as resilience, trust, security, privacy, and traceability, make it particularly suitable for Industry 4.0 applications. Nevertheless, blockchain technology exhibits limitations, such as low throughput, high resource consumption, low energy efficiency, low scalability, and considerable transaction storage delays. These drawbacks are particularly problematic in industrial environments where IIoT devices generate vast amounts of data.

Considering the aforementioned challenges, in this section the data source layer of the architecture is presented, which is tailored for lightweight, energy-efficient operations with high scalability and throughput capabilities. It encompasses both the machine level, where IIoT and control devices reside, and the production line level, typically comprising various machines. A DAG DLT at the data source level is employed, and several improvements to existing DAG DLTs in terms of storage, cryptography, and consensus are presented. Through simulations, the efficacy of the proposed optimizations is demonstrated, aiming to achieve the highest possible participation rate of IIoT devices in the DLT network.

By addressing the unique requirements and challenges of IIoT within the context of Industry 4.0, this research contributes to the development of a scalable, secure, and efficient DLT architecture that enables the seamless integration of cutting-edge technologies and fosters the growth of IIoT in various industrial sectors.

4.3.2 Improved DAG for the Data Source Layer

In order to establish the most appropriate DLT solution for this layer, a comparative study between the most promising DLT solutions for lightweight environments that currently exist is performed [222] [223]. In Table 4.1 several key characteristics and technical aspects of each solution are compared:

- **Scalability.** It is an important aspect in environments such as Industry 4.0 where the number of devices and the amount of data could grow exponentially [224].
- **Throughput.** The huge amount of data that are generated by IIoT devices [225] demands a DLT with a high capacity of transaction processing.
- **Data structure.** The data structure of a DLT is a relevant information that could reveal its storage needs and its capacity to process data.
- **Validation time.** IIoT environments demand a DLT solution that offers reasonable validation times [122].

- **Energy efficiency.** Many IIoT devices are battery-powered [226], thus they would not be able to support the load of an energy inefficient DLT.
- **Public / private.** Public and private DLTs have different characteristics, requirements and applications [227].
- **Smart contracts.** Automated and immutable digital contracts that execute on top of a DLT. They are meant to greatly improve the business process of Industry 4.0.
- **Platform languages.** The language in which a DLT is written can affect its growth possibilities, efficiency, and interoperability with other DLTs.
- **Popularity.** The "popularity" may indicate its maturity, community support and development friendliness.
- **Aimed at...** Each DLT solution is normally aimed at a specific field of application, thus it is normally adapted to the specific needs and challenges of that field.

Table 4.1 Comparison between DLTs

	Fabric	**IOTA**	**Holochain**	**Hashgraph**	**Tempo**
Scalability	Low	High	High	High	High
Throughput	High	High	Not specified	High	High
Data structure	Blockchain	DAG	Distributed data	Parallel chains	Sharding
Validation time	Seconds	Seconds	Seconds	Seconds	Seconds
Energy efficiency	High	High	High	High	High
Public / private	Private	Both	Private	Public	Public
Smart contracts	Yes	Alpha state	No	Yes	Yes
Platform languages	Go	Go, Rust	Rust	Java	Java
Popularity	High	High	Low	Low	Low
Aimed at	Consortia	IoT	Varied	Varied	Financial

Thus, according to the comparative DLT study in Table 4.1, the most appropriate DLT solution for the data source layer is the DAG type. A DAG-based DLT was chosen in order to promote scalability and fast transaction processing from the large scale IIoT. DAG type DLTs were first introduced in [228]. A DAG DLT does not use miners to validate transactions, instead, the nodes that issue a new transaction must approve two previous transactions and perform a small amount of PoW, in order to avoid spam in the network. Transactions can therefore be issued without fees, facilitating micro-transactions. DAG DLTs offer huge scalability and throughput [145], as the more transactions are issued, the faster and more secure the network is. Furthermore, the lack of

mining makes DAGs highly efficient and suitable for lightweight devices. Finally, the IOTA[1] DAG platform was specifically created for IoT devices, it supports both public and private networks and enjoys a very high level of popularity and support [229].

The proposed DAG DLT at the data source layer is designed to enhance the efficiency and reliability of data exchange in IIoT environments. This architecture comprises clusters of lightweight devices, such as sensors, thermostats, actuators, and other components typically found in industrial machines. These devices work together to gather and transmit data, enabling precise monitoring and control of various processes within the production line.

As depicted in Figure 4.1, each cluster of lightweight devices is managed by a more powerful Edge node, which serves as an integral part of the industrial control system within a production line. This Edge node is responsible for processing and analyzing the data gathered by the devices in its cluster, as well as coordinating their actions to maintain optimal performance and efficiency. In this context, the industrial control system functioning as the manager of the IIoT devices is referred to as the Cluster Head (CH).

The DAG DLT architecture allows for a decentralized and scalable approach to data management, as each CH operates independently and communicates directly with other CHs in the network. This setup ensures that the data remains secure and tamper-proof, as any attempts to manipulate the information would require compromising multiple CHs. Furthermore, the DAG structure enables faster data processing and consensus, as transactions and data updates can be executed in parallel, rather than sequentially, as in traditional blockchain systems.

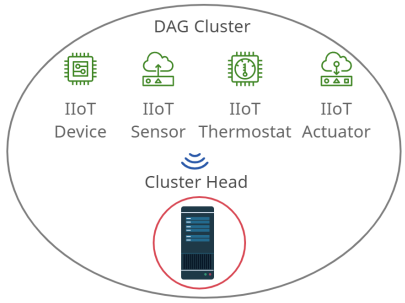

Fig. 4.1 DAG cluster within a production line

While DAG DLTs offer significant advantages for IoT focused solutions, there is still potential for further enhancements [211]. In the subsequent sub-sections, several improvements are proposed to address the limitations and optimize the performance of

[1] https://www.iota.org/

DAG DLTs in IIoT environments:

- Lightweight devices participation in the DLT. The participation of lightweight IIoT devices in the DLT network is encouraged by boosting their hardware capabilities.

- Storage burden reduction. The size of the data that has to be stored in the DAG is drastically reduced.

- Cryptography. The energy consumption and improve the performance of the cryptographic algorithms is reduced by proposing the use of more efficient solutions.

- DAG anti-spam mechanism improvement. The computational burden and energy consumption of the anti-spam mechanism of the DAG is reduced while boosting its transactions processing capabilities.

Lightweight devices participation

In this work a high degree of lightweight nodes participation within the DAG network is achieved, instead of following the usual approach where the IoT devices only send data to a more powerful node that supports all the burden of the DLT. Promoting lightweight devices participation within DLTs, especially for sensing devices is a significant challenge, as shown in other works such as [230], where authors use a software architecture specifically designed for a trust-less water management system where IoT devices can directly transact sensed data on a blockchain network. For this purpose the architecture makes use of swap memory in order to increase the storage capacity of IIoT devices.

The use of common Random Access Memory (RAM) memory would be the most optimal choice in terms of performance. Nevertheless, RAM memory is volatile, which means that it requires power to maintain the stored information. It retains its contents while powered on but when the power is interrupted, the stored data is quickly lost. Furthermore, RAM memory would entail a considerable cost of money. Therefore, for this purpose the use of Solid State Drive (SSD) swap memory is introduced. A SSD is slower than RAM memory [231], but, it has numerous relevant advantages such as: low cost, high availability on the market, and the fact that it is non-volatile. By using this approach is expected to have a greater participation of lightweight IIoT devices in the network.

Storage

In the modern Industry 4.0, the data that are generated by sensors and other IIoT devices, cloud-based solutions as well as business management are continuously increasing. According to [216], it is known that industrial data has reached a total volume of more than 1000 Exabytes annually, with a clear upward trend. The data that are generated by the lower level of a smart factory, directly from the machine tools and the human operators is of high importance for an enterprise, as these data are used and analysed in order to provide relevant information to the higher levels of the enterprise.

In blockchain and other DLTs, there typically are three main storage approaches:

1. Store all the data on the DLT. This is the most complete approach, as it assures the integrity, immutability and availability of the data.

2. Store the data in a centralized database or a decentralized database, and keep the references of the data in the DLT, in order to assure its integrity.

3. Store the data in the cloud, either a centralized cloud such as Amazon Web Services (AWS), or a blockchain-based cloud, such as Storj. This approach involves the outsourcing of the data storage.

Storing such great amounts of IIoT data in the DAG DLT is costly and inefficient, as it requires the use of high storage capacity industrial control devices and high capacity networks. Furthermore, one of the top priorities of this proposal is to encourage as much as possible the participation of the lightweight IIoT devices in the DAG network. Thus, the most optimal balance between trust and lightweight is the second storage approach. The externalization of the data storage to the cloud could also be studied, however, in this work an external consortium network that can be considered as "the cloud" is already proposed. Furthermore, in this section, data storage issue within the smart factory at a data source level is being discussed.

The use of the IPFS to solve the storage problem is proposed. With IPFS, data are immutably secured and timestamped, without having to attach all of it to the DAG DLT. IPFS is a distributed P2P file system. In this architecture, IPFS storage would be handled by the powerful Edge nodes that are defined within the DAG DLT, while IIoT devices would only store IPFS hashes. Therefore, there is no need to introduce additional devices in the architecture. It is worth mentioning that IIoT devices would be capable of storing IPFS hashes following the proposal that is made in Section 4.3.2. Using this approach the participation of lightweight devices within the DLT is encouraged while preventing storage overload. This approach contrasts with the typical blockchain

architectures for IoT where lightweight devices are practically isolated from the DLT; they only send data to more powerful nodes.

Figure 4.2 shows the flowchart of the data storage process in the DAG model. When a file is added to IPFS, the file is divided into various blocks, and all of the blocks are given a unique cryptographic hash. Then, IPFS removes the duplications that are present across the network. Each network node in IPFS stores only the content of its interest, along with some indexing information. Thus, IPFS is more suitable than a centralized database for keeping the production lines databases linked in a secure and decentralized way. Hence, in this solution, the IIoT data that are generated by each production line DAG DLT are grouped and stored in IPFS, while the DAG stores the hash of the IPFS files containing the IIoT data. The cryptographic hashes of the files can then be used to find the actual location of the file.

Consequently, IPFS presents the following issue: anyone can access the shared files if they have their hashes. Therefore, a control access mechanism such as the one in [232] should be defined. Alternately, the data could be encrypted before being added to the database. Moreover, data querying performance in IPFS could be a bottleneck, since for the read queries, IPFS requires to resolve remote nodes and download objects via the internet [233].

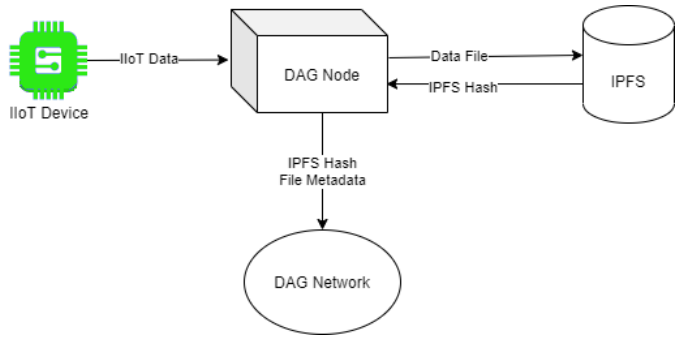

Fig. 4.2 Data storage process flowchart

Cryptography

Cryptography, an essential component of blockchain and most of the existing DLTs, serves as a fundamental pillar in ensuring the security, authenticity, and integrity of transactions and data within these systems [221]. Its role is not merely limited to encryption and decryption processes; it extends further to include digital signatures, hash functions, and zero-knowledge proofs, all of which contribute to the robustness and reliability of these technologies. Nevertheless, the implementation of cryptographic

methods involves intricate considerations of balancing security, cost, and performance [234]. Therefore, the choice of cryptographic measures should be made meticulously, with careful consideration of this trade-off, to maintain an optimal balance that ensures the secure and efficient operation of blockchain and DLT systems.

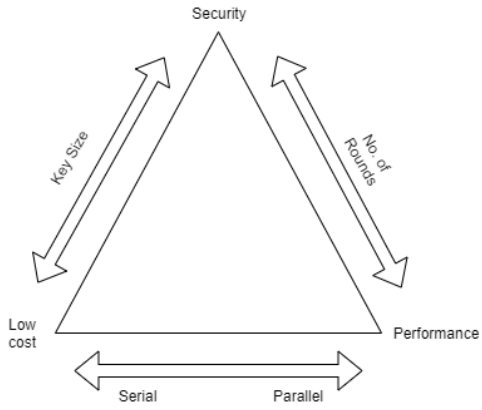

Fig. 4.3 Cryptography trade-off between security, cost and performance

A DLT usually includes two types of cryptographic functions:

- **Hash functions.** A hash is a cryptographic function that is easy to check, but difficult to forge, allowing the generation of digital signatures that users need to authenticate themselves or their transactions data. In blockchain, hash functions are also used for block linking and for mining following the PoW protocol.
- **Public-key (asymmetric) encryption digital signatures.** A digital signature scheme is commonly used for verifying the authenticity of data. Digital signatures assure that a certain message was created by a known sender, and that it has not been altered in transit. Each digital signature scheme has a pair of keys. The private key is used for signing messages and the corresponding public key for checking the signature.

In the DAG network mining is not present. However, in order to avoid spam in the network, each device is required to perform a little PoW in order to send its data to the network. This PoW process is much less expensive than Bitcoin's PoW. However, for lightweight devices, it can still be too costly. This aspect is discussed more in detail in Sect. 4.3.2, where a reputation based system is proposed in order to avoid trusted IIoT devices doing a great amount of PoW. The most popular DAG DLT, IOTA, is currently using Blake2b as its main hash algorithm. However, the aim is to further improve this aspect by using a more lightweight hash algorithm such as Quark [235], which offers

about five times more throughput that other lightweight functions such as Spongent [236], while consuming just 1.77 more μW [126]. With this proposal a performance boost is expected, as well as a significant reduction in energy consumption.

Digital signatures are used to verify the integrity of the data [237]. The first blockchain system, Bitcoin, uses the Elliptic Curve Digital Signature Algorithm (ECDSA) to generate public and private keys for its crypto-currency wallets. On the other hand, IOTA DAG DLT platform recently implemented the Edwards-curve Digital Signature Algorithm (EdDSA) [238] digital signature scheme, which was designed to be faster than the current digital signature schemes while maintaining the same level of security. Its key advantages for lightweight devices are higher performance and straightforward, secure implementations [238]. Before EdDSA, IOTA had formerly used a ternary system, which was proved to be ineffective while preventing IOTA's mass adoption in real world scenarios. However, the former ternary system had a significant advantage over EdDSA; it gave IOTA resistance to quantum attacks [211].

EdDSA is replaced with an efficient post-quantum digital signature scheme finalist from the National Institute of Standards and Technology (NIST) post quantum standardization process competition [239]. FALCON [240] has been chosen for its efficiency claims, as well as for the availability of trustworthy implementations in several programming languages. With this proposal, is expected to have a robust, quantum resistant digital signature algorithm without complex and impractical ternary architectures, while sacrificing (possibly) just a bit of performance.

Anti-Spam Mechanism

In a DAG DLT, when a node issues a new transaction, it has to solve a cryptographic hash "puzzle" similar to that of the Bitcoin blockchain, in order to avoid malicious nodes spamming the network. Subsequently, the issuer has to validate two other transactions that are in the DAG. The issuer node has to check if the two approved transactions conflict by examining the network's history, and if it discovers a conflict, it will not approve them. With this mechanism, every node that issues a transaction automatically contributes to the security of the network. The DAG network defines a transaction's cumulative weight as the sum of the weights of other nodes that directly or indirectly approved that transaction, including itself. Thus, the cumulative weight determines the transaction's importance in the network [211].

Designing a novel consensus algorithm for DAG DLT platforms such as IOTA and tackling the double spending issue for financial transactions is out of the scope. In this work the focus is on improving the most heavy and inefficient aspect of the consensus process for lightweight IIoT devices: the PoW anti-spam mechanism. In order

to achieve this, a novel reputation based mechanism that would drastically reduce the amount of PoW that IIoT devices must perform before issuing transactions in the DAG DLT is proposed. This approach is focused on reducing the computational burden and the energy consumption of the lightweight devices that participate in the DAG DLT.

First, there is the need to define what a "valid transaction" is in IIoT. In financial transactions this concept is relatively simple, as transactions only contain numerical cryptocurrency operations and they can be easily validated by checking the signature and the inputs and outputs. According to the present industrial experience, in an IIoT scenario a range of expected data values from the IIoT devices must be defined. Kuemper et al. [241] deeply analyzed this topic and proposed a framework that utilizes a metadata annotation for IoT data attributes supporting structures related to quality metrics. The framework provides mechanisms for data attribute representation and comparison including a model-based quality analysis for IoT data sources. Furthermore, adding suspicious data values to a side chain for manual revision as an additional measure is also proposed.

A Reputation (R) score that goes between zero and ten (0-10) is defined. For each IIoT device, R would be calculated by the control Edge nodes in a defined time interval called Validation Period (VP), similar to other work where a reputation system is proposed [89]. A reputation of zero would imply a total impossibility of participating in the network. Each IIoT device starts with a medium reputation of five during its first VP. This condition is defined due to the fact that a new device would not have any history of transactions at the beginning, thus the calculation of R would not be possible. Then, the reputation score would increase or decrease based on the behaviour of the device. The aforementioned score is inversely proportional to the amount of PoW that a device has to perform (i.e. the more reputation score, the less PoW is required). The behaviour of a device is defined by **the number of Validated (V) transactions** that it had in the past, as well as **the total amount of Transactions (T)** of that device. Thus, network spamming can effectively avoided, while reducing the computational burden for IIoT devices. The proposed reputation system calculation can be represented as follows:

$$R = \frac{V}{T} \times 10 \tag{4.1}$$

Below a practical case scenario of a new honest device that would join the DAG network using the proposed reputation mechanism is defined. In Section 4.7.1 a case scenario where a device has been compromised is considered, in order to analyze the security of the reputation mechanism.

New honest device joins the network scenario

When a new device D_1 joins the network, it starts with an initial reputation score R of 5, and both its total number of transactions (T) and validated transactions (V) are equal to 0. As the device begins to emit transactions, its T and V values change accordingly. Let's consider a scenario where D_1 has already emitted 25 valid transactions, but only 15 of them have been referenced and validated within the DAG at the moment of the reputation calculation. In this case, T and V would be updated to 25 and 15, respectively.

The reputation score R is then recalculated, taking into account the updated T and V values. The higher the proportion of validated transactions to the total number of transactions, the more the reputation score R will increase. In this example, the updated R value would be 6, which results in the honest device having to perform a lower amount of PoW.

As time progresses and more transactions are validated within the DAG, the reputation score R will continue to increase, eventually reaching a maximum value of 10. This continuous increase in R reflects the growing trust in the device's honest behavior and leads to a reduced PoW requirement. As a result, the device's energy consumption and computational burden are lowered, making the network more efficient and sustainable.

4.4 Layer 2: Bridge layer

4.4.1 Introduction

So far, in the presented scenario (depicted in Figure 4.4), the data generated by IIoT devices is securely gathered and stored by the DAG DLTs and IPFS storage. However, Industry 4.0 does not stop at the machine data level, and this data that is being gathered at "the lowest level" needs to be exploited and processed by "higher" levels to derive and build actual information, such as machine and IIoT fleet status, machines predictive maintenance (by AI algorithms), compute overall process productivity, etc. Hence, these "upper" processes **need to access and process heterogeneous data from all cluster plants**. However, accessing and processing all the raw machine data from all production-level DAG-type DLTs is not a straightforward procedure, essentially due to the:

- Heterogeneous machine data. Data could be expressed in different units of measure depending on the machine provider, machine version, country, etc. They could have a distinct number of decimal places, obey different standards, or they can include certain errors or variations. This problem stems from the fact that

according to Jirko et al. [242], *"machines within a complex system are produced by different manufacturers with different data models and interfaces"*. Consequently, this issue affects industrial interoperability and integration, thus, creating a detrimental impact on the ability to effectively process data using disruptive technologies, such as Big Data or AI.

- Lack of efficiency and security when accessing machine data [243]. It is not efficient nor secure to directly delegate the responsibility to external data exploitation services to access and process the raw machine data into "readable" plant-level data. Accessing machine data means that each data exploitation service needs to be a client of every production line DLT that wants to access data from. Additionally, these services would need to simultaneously process all the data from all machines and homogenize it accordingly. This approach lacks efficiency as the data exploitation services would spend a high amount of time accessing and homogenizing data before exploiting it. This is not utterly secure either since it breaks the data custody chain and mixes responsibilities, as in each data exploitation service, the actual data format being used to be exploited becomes obscure, and the traceability and integrity are compromised.

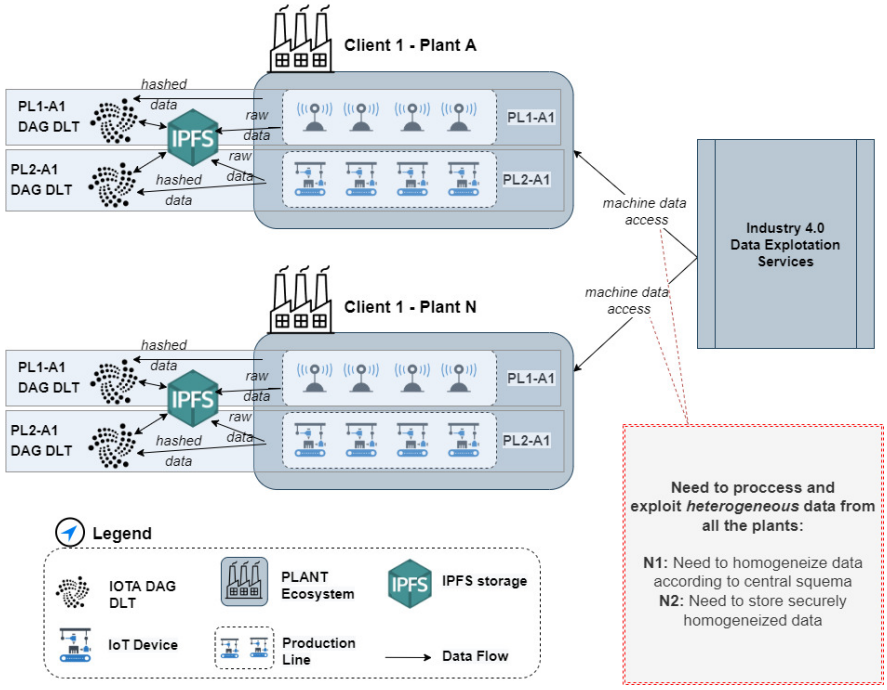

Fig. 4.4 Industry 4.0 motivating Scenario.

Furthermore, such complex ecosystems require proper monitoring to achieve higher efficiency rates by notifying human operators of probable performance gaps and possible disruptions through the presentation of data. Even though, theoretically, the use of DLTs improves the security of industrial processes, many attacks, such as Distributed Denial of Service (DDoS) attacks, are still possible. Thus, they need to be identified and mitigated as soon as possible to avoid the disruption of industrial production. In addition, proper monitoring can also help mitigate errors, and optimize production processes and associated costs, which are known to be critical in industry [244].

In this context, this second layer of the multi-DLT architecture presented in this thesis aims to mitigate these problems, by providing means to that:

- The raw machine data that resides in DAG-type DLTs are securely and consistently homogenized by a secured and traceable process. This will ensure that the data conforms to a common data model, thus, providing interoperability so that processes at higher levels can exploit the data in a consistent manner.

- The homogenized data is securely stored and accessed, ensuring its integrity and availability. This will ensure trust in the data throughout the whole process, from where the data is generated from the production lines to where it is exploited and processed at a higher level. Processing raw IIoT data through a DAG DLT is a pointless approach if, at a higher level, there is a centralized and non-persistent data structure where the data can be easily tampered with [245].

- The whole industrial architecture must be carefully monitored using a monitoring system that is able to analyze all components securely; the IIoT sensors and actuators, the DLTs, the storage systems, etc. This analysis is required for performance and security optimizations and prevention to avoid the malfunction of critical processes.

Consequently, the "Bridge Layer" section presents the following contributions to the described issues:

1. A "data homogenization" process for solving data interoperability issues that relies on the use of decentralized blockchain oracles as a trustworthy source for the target data model scheme the data needs to conform to. An oracle architecture is crafted by employing a more versatile blockchain platform to improve simplicity and provide more interoperability capabilities. Finally, the resulting homogenized data is stored in a blockchain-based solution for trustworthy access and processing.

2. A prototype that implements the secure data homogenization process that: (i) accesses raw machine data stored in DAG DLTs, (ii) gets the target data model schema from the oracles, (iii) performs the data homogenization from the source data scheme to the target data schema, and (iv) stores the homogenized data into a "plant level" blockchain network so that it can be consistently accessed and processed by other services. The monitoring system of the aforementioned scheme is also implemented.

3. A monitoring system for the proposed scheme to track the quality of the retrieved data, the performance of the network, the usage of each oracle, billing reports, security incidents, etc. A monitoring architecture Application Programming Interface (API) for data retrieval is implemented and visualizes using the ELK (Elasticsearch, Logstash and Kibana)[2] stack.

4.4.2 Interoperable Plant Blockchain for Homogenized Data via Smart Oracles

In this section, the proposed solution for machine data interoperability and trustworthy storage of plant-level data is described. First, the proposed data homogenization process using smart oracles is described, and then the design of a monitoring scheme for the proposed architecture is presented.

Figure 4.5 depicts the proposed solution, in orange, on top of the motivating industrial scenario that was presented above. Specifically, in grey, we have N smart factories where the IIoT data is processed using DAG DLTs along with IPFS decentralized storage. Additionally, in orange, we have the proposed extension that is addressed in this work. A data homogenization service that makes use of blockchain oracles and has the resulting data stored in an interoperable external blockchain was added. On top of the scheme, we also have a monitoring system for the whole architecture.

Data Homogenization via Decentralized Oracles

As mentioned in the motivating scenario, the actual IIoT data is stored inside an IPFS storage system, while the data-source DAGs would only store the hashes to reduce the storage burden of the DLTs. In the proposed scheme, after receiving and storing the raw IIoT data hashes from IPFS, a data homogenization service that is executed periodically would make a call to an external decentralized oracle service to retrieve the data model used for the data homogenization process. Blockchain oracles are needed since smart contracts are unable to access external data sources in a trustworthy manner. Hereafter,

[2] https://www.elastic.co/

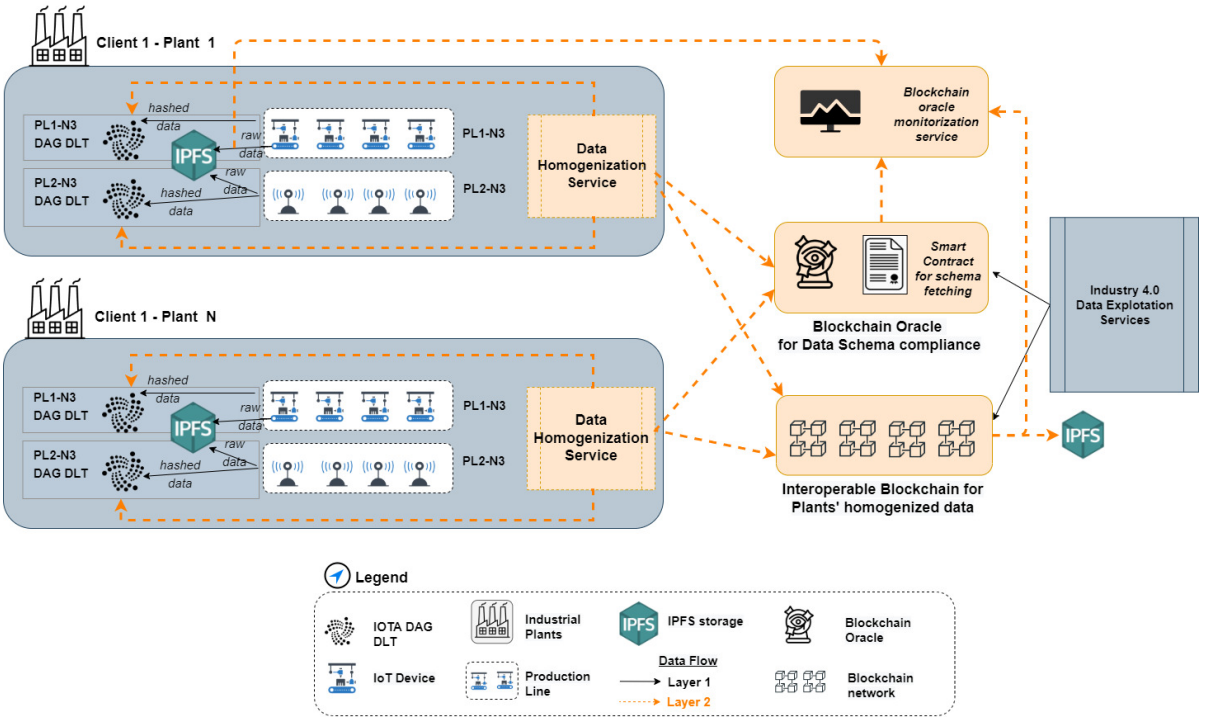

Fig. 4.5 The proposed interoperable plant blockchain and data homogenization via decentralized Oracles scheme.

once the data model is received from the oracle, the data homogenization process starts its execution. The homogenization process consists of converting raw IIoT data into a standardized data scheme according to the given data model. Finally, the data homogenization service would then send the homogenized data to an interoperable plant blockchain, which in turn stores it inside the IPFS storage system and keeps its references within the immutable ledger. Figure 4.6 depicts the sequence diagram of the presented homogenization process.

Therefore, the main purpose of the interoperable plant blockchain is to store and manage the smart plant securely homogenized data references and provide access control to IPFS. This blockchain would also unify the data management of different industrial plants belonging to the same business conglomerate. Finally, this ledger would act as a bridge between the DAG DLTs that process the data from IIoT devices inside production lines, and other hypothetical DLT connections with other organizations within a hypothetical decentralized business consortium network.

Consequently, interoperability capabilities are required at this level. To connect the

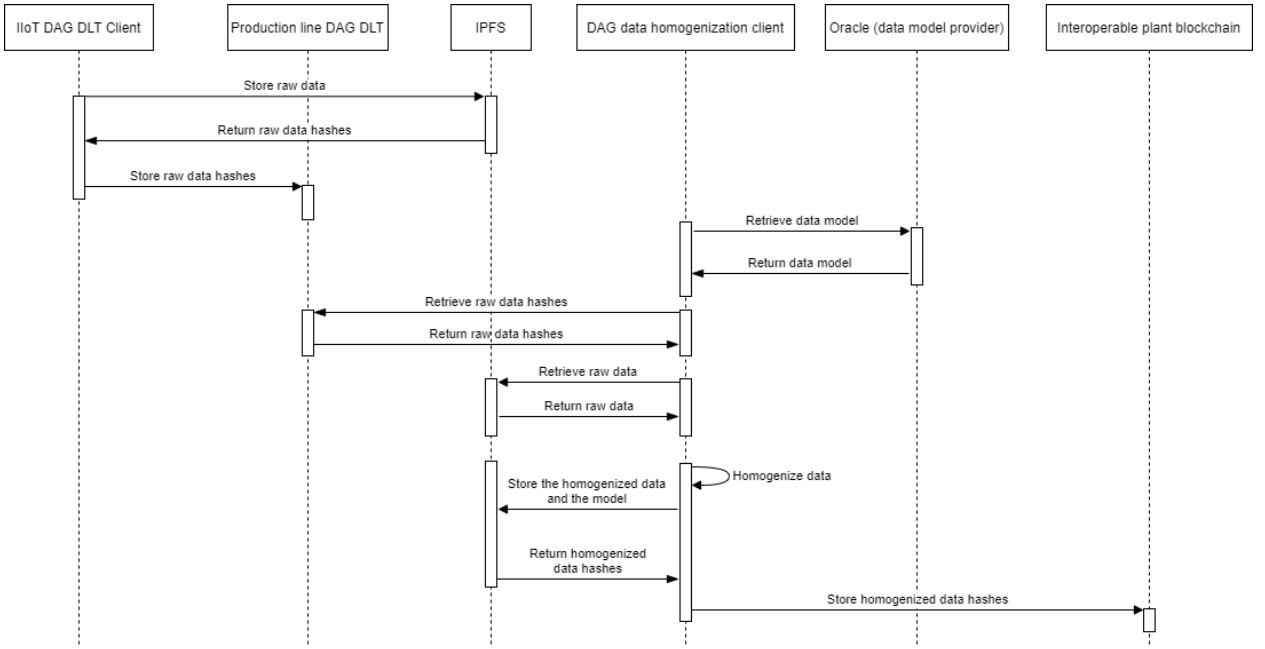

Fig. 4.6 Sequence diagram of the proposed oracle-based architecture

production lines DAGs and the plant blockchain, a smart contract-based notary scheme that interacts with a smart contract from the destination blockchain is implemented to transfer the data securely.

Application Example: One real-world example application of the approach described could be a system for collecting and storing data from sensors in an industrial plant. In this system, the raw data from the sensors would be stored in IPFS, and the hashes of this data would be recorded in a DAG DLT. The data homogenization service would periodically retrieve a data model from a decentralized oracle service, and use this model to convert the raw sensor data into a standardized format. The homogenized data would then be stored in IPFS and recorded in the interoperable plant blockchain.

This system could be used to ensure the integrity and traceability of the sensor data, as the data would be stored in a decentralized and immutable manner. It could also help to facilitate data interoperability, as the standardized data format would make it easier for different systems and applications to make use of the data. Additionally, the use of oracles to retrieve the data model from an external source could allow the data homogenization process to be updated and improved over time, as the oracle could provide access to the most recent data model. Finally, this approach also enables the data homogenization process to be updated and improved over time.

Monitoring System Architecture

The purpose of the proposed monitoring system is to visualize and analyze the industrial data throughout the whole process, since it is generated at an IIoT level up until it is homogenized and exploited at a plant level, along with all the elements that intervene in the aforementioned process. These elements go from the IIoT devices to the DLTs, and IPFS storage until the blockchain oracles. A monitoring scheme covering all the elements apart from the IIoT devices is required to check the quality and integrity of the retrieved data, the status and usage of each element, accrued financial costs, and other financial information for future business-related use cases. Furthermore, in modern Industry 4.0, strict monitoring is also required so cyber-attacks or performance issues can be rapidly identified and mitigated. For example, monitoring the number of active devices, their effectiveness, or temperature can provide a holistic picture of the overall productivity and weaknesses of the plant. Monitoring of Information Technology (IT) elements could help us identify performance bottlenecks, vulnerabilities, and cyber-attacks, and optimize the IT infrastructure associated costs [246].

To make the monitoring system as efficient and secure as possible, three guidelines have been followed when designing it [247]: (i) the collection of metrics should not have a significant impact on the performance of the employed DLTs or on the data homogenization process, nor it should create a massive data traffic overhead; (ii) it should be as modular as possible to support different DLTs and oracle services; and lastly, (iii) the defined metrics should be defined to cover multiple industrial scenarios.

The proposed monitoring system consists of five modules (Figure 4.7): (1) IIoT data monitoring agent; (2) storage monitoring agent; (3) oracles monitoring agent; (4) the DLTs monitoring agent; and (5) the monitoring system core.

The IIoT Data Monitoring Agent handles data collection from numerous industrial IoT devices, ensuring the reliability and accuracy of this information. This processed data is then securely stored and managed by the Storage Monitoring Agent, which also facilitates data backup and transaction logging.

Acting as an intermediary, the Oracles Monitoring Agent verifies the authenticity of external data before it is incorporated into the blockchain. Simultaneously, it monitors the performance of the oracles to maintain the system's overall data accuracy.

The DLT Monitoring Agent oversees the performance and security of the underlying blockchain, checking transactions and identifying potential anomalies.

Finally, all these modules are managed and coordinated by the Monitoring System Core. This core module processes and analyzes data, guides decisions, and provides an interface for administrators to oversee the entire system.

Effective monitoring requires strategic placement of measurement probes, without

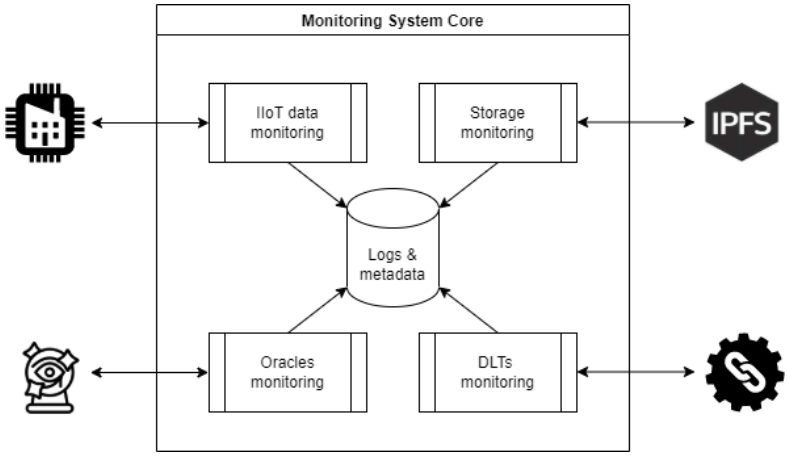

Fig. 4.7 Monitoring system architecture.

affecting in any manner the flow of the data and thus causing more latency and overall poorer performance. Furthermore, the monitoring system must be designed in such a way so the data cannot be fraudulently accessed and tampered with through it. Consequently, similarly to other works such as [248], cheap lightweight Field Programmable Gate Array (FPGA) devices with limited access to the actual data for the monitoring tasks have been used. Thus, apart from avoiding illegal access to the data, using cheap devices avoids a significant increase in the operating costs of the architecture. Figure 4.8 shows the monitoring probes placement process across the presented architecture.

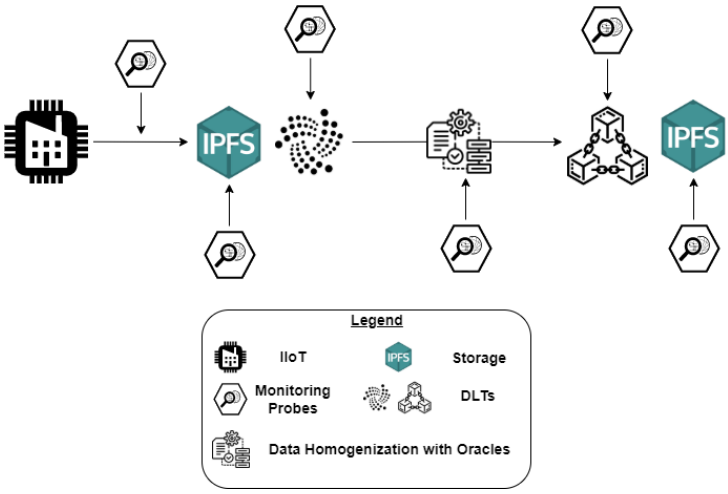

Fig. 4.8 Monitoring probes placement across the presented process.

The monitoring process is composed of the following four steps:

1. First, there is the need to place a monitoring probe at the IIoT level so the original raw data can be monitored at the exact source before being stored or processed by any other agent.

2. The second step is to monitor the data when it arrives at the IPFS-DAG tandem. The comparison between the data that comes from the IIoT devices with the data that is finally stored and processed in IPFS and the DAG can help identify possible man-in-the-middle and DDoS attacks or mere transmission failures. Apart from monitoring the data, performance and other status data from the IPFS and DAG structures can also be monitored.

3. The third step is to monitor the data homogenization process, along with the employed oracles, so we can ensure that the process has been correctly executed. Regarding the oracle scheme, we can comprehensively examine the usage of the oracles and possible incurred costs, as well as possible performance and security issues.

4. Finally, the last probes monitor the homogenized data at the plant level structures; the interoperable plant blockchain, and the related IPFS partition. Monitoring this part of the architecture helps us ensure that the homogenized data has been correctly stored and processed. We also need to make sure that there are no performance or security issues that can compromise the data prior to exploitation for business processes.

4.5 Industry 4.0 Business-oriented Blockchain Decision Tree

4.5.1 Introduction

In the realm of Industry 4.0, blockchain technology harbors the potential to revolutionize business operations by enhancing transparency, efficiency, and security. However, the creation of an appropriate blockchain solution that meets the unique needs of various industrial sectors demands a specific and targeted approach [249]. Blockchains, specifically those catering to business needs, are heavily influenced by the demands of their respective sectors. Therefore, their development necessitates a specialized approach.

The motivation for generating a decision tree to architect a fit-for-purpose blockchain for the Industry 4.0 business environment arises from the disparate needs of various industries in terms of their data management frameworks. Industries such as supply chain

management might necessitate a blockchain capable of tracing goods from manufacturing to delivery, while sectors like manufacturing may need a blockchain that seamlessly integrates with their current ERP systems. Furthermore, many scenarios could demand the execution of automated agreements (smart contracts) over a blockchain. Such contracts might often require external data, which is a complex process to achieve [24]. Thus, a one-size-fits-all blockchain solution might not cater to the needs of all sectors.

Given the layered architecture of the proposed model, it's essential to clarify why the focus is solely on the third layer - the business consortium layer. The data source layer and plant-bridge layer, while fundamental, serve different purposes. The data source layer collects and manages the raw data, and the plant-bridge layer acts as a conduit for transferring this data. However, the business consortium layer is where the data is transformed into actionable insights and decisions. It is at this layer that the blockchain technology finds its most crucial application: facilitating secure, transparent, and efficient business transactions. Therefore, tailoring a blockchain solution that addresses the unique requirements of this layer is a pressing concern.

In this section, the application of blockchain in the business-oriented sector of Industry 4.0 is studied along with the requirements of this particular field. Some existing blockchain solutions for Industry 4.0 and their limitations are also considered. Lastly, a decision tree for devising a suitable blockchain solution for the Industry 4.0 business sector is proposed. It takes into account its distinct needs and prerequisites. An extensive analysis of the characteristics a business-oriented blockchain should possess is undertaken, based on several input parameters pertinent to the given needs.

4.5.2 Industry 4.0 Business Requirements

In this section, the main requirements of a business-oriented Industry 4.0 environment and the characteristics that it should possess to meet the needs of this field are enumerated. Thus, 11 fundamental requirements based on the current knowledge on the field [250] [251] are defined:

1. **Security.** One of the most critical requirements for a business Industry 4.0 environment is security. Due to the integration of various technologies and the management of sensitive information, there is a significant risk of cyber-attacks and data breaches. Security can be achieved by implementing strong access controls, data encryption, network monitoring and trustworthy traceability.

2. **Transparency.** Another key requirement for a business Industry 4.0 environment is transparency. Typically, it is challenging to keep track of data and processes due the integration of various systems and devices. Therefore, businesses must

ensure that their systems provide traceability to enable better decision-making and transparency.

3. **Data Privacy.** Data privacy is a crucial requirement in the Industry 4.0 and most digital fields. Nowadays, businesses collect a great amount of data from various sources. Therefore, there must be an assurance that data is collected, processed, and stored in compliance with data privacy regulations and standards in order to provide trust.

4. **Interoperability.** The Industry 4.0 environment is highly complex, with multiple systems and devices communicating with each other. To ensure seamless communication, businesses need to ensure interoperability between different systems. This requires the adoption of standard protocols and interfaces that enable communication between different systems and devices.

5. **Performance and scalability.** A successful Industry 4.0 environment requires both scalability and performance. As businesses grow and the amount of data and traffic increases, they need systems and networks that can handle the load and perform tasks quickly and accurately.

6. **Reliability.** In Industry 4.0, businesses rely on technology to run their operations. Thus, it is necessary to ensure that the systems and devices used are reliable and available when needed. This requires implementing redundant systems and devices and conducting maintenance to prevent downtime.

7. **Governance.** Governance is a critical aspect of any digital environment. The implemented systems must have a well-defined governance model that ensures the integrity, accountability, and transparency.

8. **Compliance.** Industry 4.0 companies are subject to many regulations and standards, and their systems must be compliant with these regulations.

9. **Automation.** Business process automation involves automating repetitive tasks and workflows, allowing employees to focus on higher-level tasks that require human decision-making and creativity. This approach can be applied to various business processes, including manufacturing planning, supply chain management, customer service, and financial management.

10. **Flexibility.** Finally, businesses in the Industry 4.0 environment need to be flexible so they can adapt to changing market conditions and technologies. This requires the adoption of methods that enable rapid modification, prototyping, implementation and testing of current or new technologies.

11. **Costs.** Implementing Industry 4.0 requirements requires significant investment in technology infrastructure, talent acquisition, data analytics, cybersecurity, and equipment and machinery. Companies need to carefully evaluate the costs and benefits of Industry 4.0 before implementing it.

4.5.3 Proposed Decision Tree

In this subsection, a decision tree for designing a blockchain that effectively addresses all the Industry 4.0 business requirements identified through an extensive review of existing literature is presented. The proposed decision tree comprises a series of strategically ordered questions aimed at guiding the design process towards implementing specific technical components and mechanisms that meet the blockchain requirements.

The election of the requirements and their order within the decision tree is meticulously curated based on several criteria, including their relative importance, potential dependencies, and most significantly, the logical flow of the decision tree, as shown in [252]. By organizing the aspects in a manner that adheres to established blockchain designing patterns, a more coherent and structured approach to the design process is ensured. This logical progression not only allows for seamless navigation through the decision tree but also facilitates a comprehensive understanding of the relationships between various design elements and their impact on the overall blockchain architecture.

The decision tree ensures that the blockchain system is functional, efficient, compatible with other systems, secure and maintainable, as stated in the ISO 25010 standard.

The decision tree is described in two formats: text and visual diagram format. The visual diagram is shown in Figure 4.9 and also externally[3].

1. **Security:**

   - Q1: Does the blockchain require secure access mechanisms and trustworthy traceability via identity check?
     - Yes: Consider implementing a permissioned blockchain with identity management and encryption mechanisms to ensure secure transactions and prevent unauthorized access.
     - No: Skip to Q2.
   - Q2: Does the blockchain require traceability for long periods of time?
     - Yes: Consider implementing a blockchain that does not remove old transactions.
     - No: Skip to Q3.

[3] `https://tinyurl.com/blockchaindecisiontree`

- Q3: Does the blockchain require regular security audits and updates to ensure the ongoing protection of data against emerging threats and vulnerabilities?
  - Yes: Consider implementing a blockchain that is under continuous development and based on a widely used programming language that includes the possibility of performing vulnerability scanning.
  - No: Skip to requirement 2 - Transparency.

2. **Transparency:**

   - Q1: Is full transparency of all transactions required on the blockchain?
     - Yes: Consider implementing a public blockchain that allows any user to view the ledger and its transactions.
     - No: Consider implementing a private-permissioned blockchain that provides selective access to the ledger and transactions.

3. **Data privacy:**

   - *Answer to this question only if full transparency (requirement 2) is **not** required.*
   - Q1: Is it necessary for the blockchain to provide confidentiality for some transactions?
     - Yes: Consider implementing a private-permissioned blockchain with encryption mechanisms or private channels to ensure confidentiality and selective access to transactions.
     - No: Skip to requirement 4 - Interoperability.

4. **Interoperability:**

   - Q1: Is it essential for the blockchain to interact with other blockchains or other distributed systems?
     - Yes: If the number of distinct blockchains to interoperate is significant (i.e., more than 3), consider implementing an interoperable blockchain with interoperability protocols like cross-chain atomic swaps or sidechains. If the number of blockchains to interoperate is 3 or less, consider implementing specific blockchain interoperability connectors.
     - No: Skip to requirement 5 - Performance and scalability.

5. **Performance and scalability:**

- Q1: Is it expected that the number of transactions and participants on the blockchain to grow significantly over time?
  - Yes: Consider implementing a sharded or layered blockchain architecture, off-chain processing mechanisms like state channels or IPFS storage, or use a consensus mechanism with high throughput and scalability capacities.
  - No: Skip to requirement 6 - Reliability.

6. **Reliability:**

- Q1: Is it essential for the blockchain nodes to be always online and have minimal downtime?
  - Yes: If the chosen blockchain is private-permissioned, consider implementing a Byzantine Fault Tolerance (BFT) consensus mechanism, and redundant nodes to ensure high availability and reliability. If it is a public blockchain, no further action should be needed.
  - No: Skip to requirement 7 - Governance.

7. **Governance:**

- Q1: Is a governance model to manage the evolution of the blockchain needed? Specifically, is the blockchain expected to change its characteristics at some point in time?
  - Yes: Consider implementing a blockchain with a formal governance structure such as a Decentralized Autonomous Organization (DAO) or a voting-based decision-making process.
  - No: Skip to requirement 8 - Compliance.

8. **Compliance:**

- Q1: Does the blockchain need to comply with regulatory requirements?
  - Yes: Consider implementing compliance mechanisms like regulatory reporting, identity verification, or Anti Money Laundering (AML) and Know Your Customer (KYC) procedures.
  - No: Skip to requirement 9 - Automation.

9. **Automation:**

- Q1: Is there a need for automatic enforcement or execution of business logic based on predetermined conditions? OR Is there a need for automation of tasks, such as triggering events or notifications based on specific conditions?

  - Yes: Consider implementing a blockchain with smart contracts.
  - No: Skip to requirement 11 - Costs.

10. **Flexibility:**

    - *Answer to these questions only if the elected blockchain includes smart contracts (requirement 9).*
    - Q1: Does the smart contract require access to real-world data or events that are not natively available on the blockchain? OR Does the smart contract need to be able to communicate with off-chain API?
      - Yes: Consider implementing a blockchain with a flexible smart contract platform that has the capability to access external data via oracle mechanisms.
      - No: Skip to requirement 11 - Costs.

11. **Costs:**

    - Q1: What specific additional costs could potentially arise from using the blockchain network for transactions? Could these include elements like gas fees or transaction fees?
      - Yes: Skip to Q2.
      - No: END - No further action is needed.
    - Q2: Have you analyzed the potential cost savings associated with using blockchain technology for supply chain management, such as reducing the need for intermediaries and improving traceability and transparency?
      - Yes: If the costs associated with blockchain are acceptable compared to the expected savings, no further actions need to be taken. If the costs are not acceptable, consider using a fee-less blockchain.
      - No: Skip to Q3.
    - Q3: Does your business require a high volume of transactions, and if so, will the cost of using the blockchain network for these transactions be feasible?
      - Yes: Consider implementing a blockchain with zero or near zero associated costs.
      - No: END - No further action is needed.

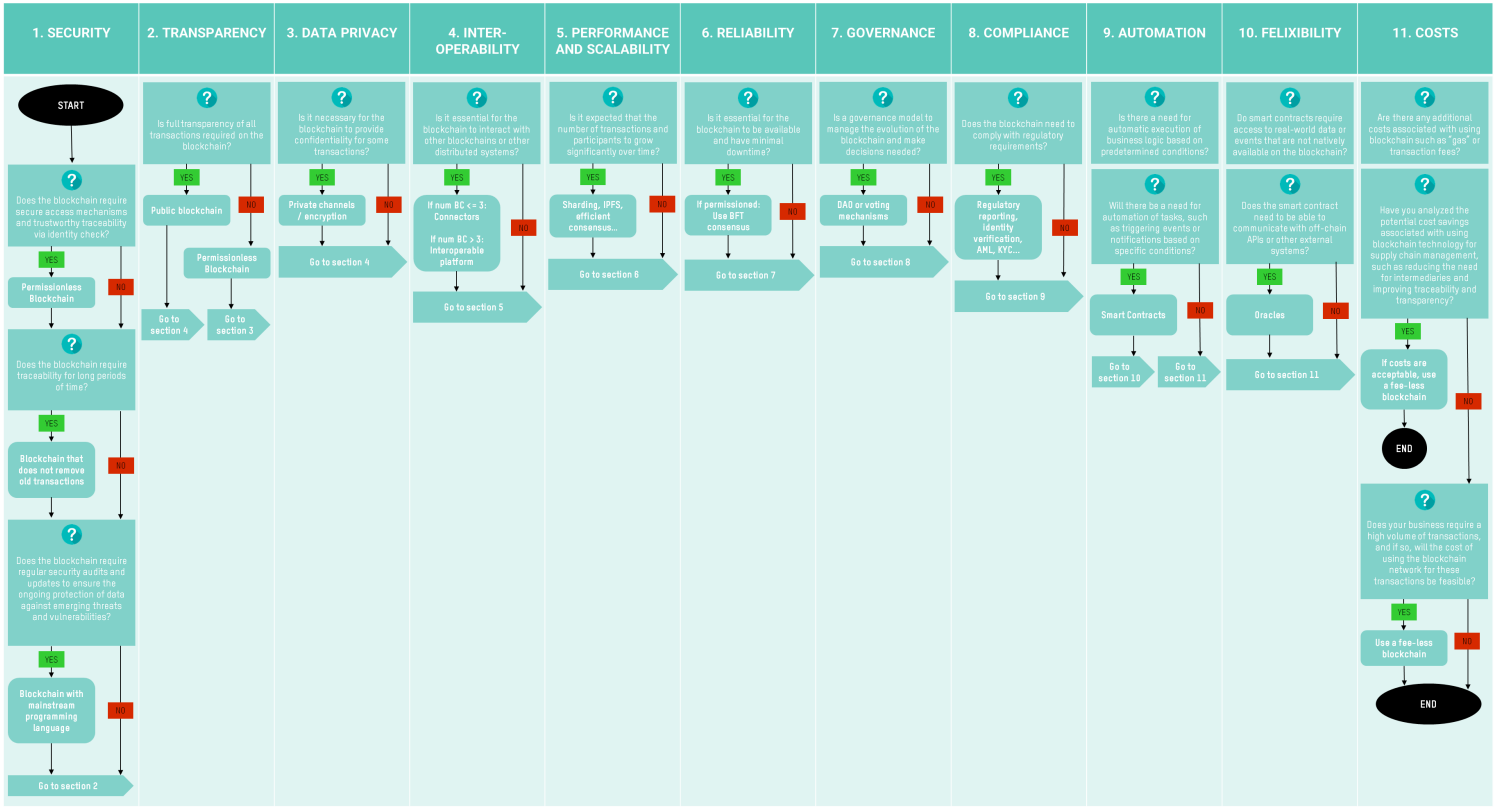

Fig. 4.9 Decision tree diagram

Blockchain platform election

In the context of the proposed decision tree for blockchain design, various blockchain platforms that align with the defined choices are identified. These platforms are classified based on the type of blockchain as per Almeshal et al. categorization [253]: private-permissioned, public-permissioned, and public-permissionless. This classification system serves as a guiding tool for selecting the appropriate real-world platform.

1. **Private-Permissioned:**

    - Hyperledger Fabric or Sawtooth with private channels, BFT consensus, regulatory mechanisms, smart contracts, and blockchain oracles.
    - R3 Corda with private channels, interoperability, BFT-based consensus, regulatory mechanisms, and smart contracts.
    - IOTA with fee-less transactions, scalable and lightweight architecture, and support for IoT use cases.
    - Private Ethereum with private channels, interoperability, BFT consensus, DAO/voting system, AML / KYC, smart contracts and oracles.
    - Quorum with private channels, interoperability, BFT-based consensus, DAO / voting system, and smart contracts.

2. **Public-Permissioned:**

    - Permissioned Ethereum with private channels, interoperability, BFT consensus, DAO / voting system, AML / KYC, smart contracts, and oracles.
    - Hyperledger Fabric or Sawtooth with private channels, BFT-based consensus, regulatory mechanisms, smart contracts, and blockchain oracles.
    - Corda Enterprise with private channels, interoperability, BFT-based consensus, regulatory mechanisms, and smart contracts.
    - Hedera Hashgraph with private channels, interoperability, BFT-based consensus, and smart contracts.
    - Ripple with private channels, regulatory systems, and smart contracts.
    - Cosmos with interoperability, fee-less transactions, and smart contracts.

3. **Public-Permissionless:**

    - IOTA with fee-less transactions, scalable and lightweight architecture, and support for IoT use cases.

- Ethereum with smart contracts and blockchain oracles.
- Polkadot with interoperability, private channels, and smart contracts.
- Solana with fee-less transactions, smart contracts, and oracles.

4.5.4 Example Use Case: Product manufacturing traceability

In this section, an illustrative Industry 4.0 use case is ellaborated, demonstrating the practical application of the proposed decision tree. This use case, though hypothetical, is constructed to mirror real-world scenarios, thereby providing a better understanding of the potential implications and benefits of implementing such a decision-making model.

Example use case

A manufacturing company called "X" makes a product that contains parts from multiple suppliers. The company wants to use a blockchain to track the movement of the aforementioned parts from the suppliers to the manufacturing facility. Whenever a part passes from one party to another, a new block is added to the blockchain. Each block includes information about the transaction, such as: time, date, location or the involved parties. Once the product is manufactured, it is shipped to a logistics company for distribution, allowing the involved parties to track the product in real time. Finally, the retailer receives the product and verifies its authenticity and integrity via the blockchain. Any issues, such as missing or damaged parts, can be traced back to their source, allowing the relevant parties to take appropriate corrective measures.

Decision tree application

1. **Security:** Yes (Q1, Q2 and Q3). Data security and long time traceability are very important, and the blockchain participants need to be identified. Thus, a permissioned blockchain must be implemented.

2. **Transparency:** No. In this case full transparency is not required, since the traceability of the parts needs to be tracked only by specific actors. Therefore, a private blockchain is the most optimal choice.

3. **Data privacy:** No. In this case no specific privacy requirement is specified. No encryption or private channels mechanisms are needed.

4. **Interoperability:** No. In this case no relevant blockchain interoperability capacity is required.

5. **Scalability:** Yes. In this case we have a product that has many parts, and these parts belong to several suppliers. While it is not clear whether the number of transactions and participants is significant, this use case requires a margin to be left in case the number of parts and suppliers increases.

6. **Reliability:** Yes. Reliability is highly recommended in order to guarantee correct data registering throughout the whole process and achieve the intended data traceability capacity. Thus, a BFT consensus is recommended.

7. **Governance:** No. In this case there is no evidence that the participants need to posses the capacity of managing the blockchain model.

8. **Compliance:** No. In this case no compliance model is needed, since the stored data belongs to industrial machinery parts.

9. **Automation:** Yes. In this case, smart contracts could be used to automatically update the blockchain with each transaction and verify that the product is complete and in good condition at each stage of the supply chain. It could also trigger alerts and notifications to relevant parties if any issues are detected, such as a missing part or a damaged product.

10. **Flexibility:** No. In this case, if the goal is simply to track the movement of the product and ensure its integrity, a basic blockchain with standard features is sufficient. Therefore, no extra features such as oracles are needed.

11. **Costs:** No (Q1). In this case, no additional costs such as gas fees, network fees, or transaction fees are present.

Therefore, according to the answers that are shown above, we need a blockchain that is: private, permissioned, scalable, reliable and with smart contracts capacity. Given these features, we can choose from several blockchain platforms: Hyperledger Fabric or Sawtooth, R3 Corda or Quorum.

4.6 Layer 3: Business layer

4.6.1 Introduction

Within the framework of Industry 4.0, blockchain can be employed to monitor the flow of products and resources across the supply chain, facilitating enhanced transparency and efficiency. Furthermore, as shown in the previous sections, blockchain and DAG

DLTs can also be used to guarantee the integrity of the data throughout the whole process, from when the data is generated from the IIoT devices until it is homogenized and shared across several plants. This approach provides robust protection, especially against data tampering attacks, while also providing transparency, immutability, accountability and interoperability to industrial processes. In addition, auxiliary tools such as IPFS are used in order to reduce the storage burden of the DLTs by keeping the actual data in IPFS and only the IPFS references within the DLTs. Thus, this approach is highly efficient and enough to guarantee industrial data integrity.

However, processing and storing plant homogenized data does not imply the end of the cycle. Generated data must be exploited so enterprises can correctly manage and improve industrial processes and generate value [24]. Industry 4.0 companies are expected to interact with a great variety of external players, including suppliers, customers, and partners. This can include other businesses, government agencies, energy and materials providers, research institutions, and other industry groups. In particular, Industry 4.0 enterprises often engage in collaborations and partnerships to access new technologies, shared data and analytics, and required resources for their processes. Additionally, Industry 4.0 enterprises often rely on external service providers for specialized expertise in areas such as data analytics or cybersecurity.

Currently, the management of business processes in the industry is predominantly centralized, unverifiable, untrustworthy, and lacking automation [8]. This is where the notion of smart contracts emerges. Smart contracts are self-executing agreements with the conditions of arrangements between two or more parties embedded directly into lines of code. Within the scope of Industry 4.0, smart contracts hold the potential to address several pertinent issues, such as inadequate automation, traceability, or data manipulation [254]. By automating the completion and monitoring of transactions, smart contracts deliver a high level of automation, which helps to prevent time loss and human-induced errors. Smart contracts also contribute to enhancing transaction transparency and security [27]. As they are self-executing and transparent, disputing the terms of a smart contract or altering it without other parties' consent becomes challenging. The trust among parties and reduces fraud risks. Another issue that smart contracts resolve in Industry 4.0 is regulatory compliance [255]. By streamlining the completion and tracking of transactions, smart contracts can assist businesses in adhering to relevant regulations and standards.

Nonetheless, smart contract business processes adoption by Industry 4.0 enterprises is still uncommon due to several challenges that still need to be solved [24] [256] [257]:

- Lack of interoperability. It is difficult for different organizations to interoperate since there are many different smart contract platforms and distinct smart contact

data structures.

- Scalability and performance limitations. Smart contracts can be heavy on blockchain networks, which can slow down the performance in high-transaction situations.
- Limited data privacy. Smart contracts are executed on a distributed ledger, which means that all of the data used in the contract is public and transparent. This can be a problem for organizations that handle sensitive or confidential information.
- Limited access to external data. Smart contracts can only work with data that is stored on the ledger on top of which they are written.
- Fees. Some smart contract-compatible blockchains require fees for each transaction. This may incur high costs for enterprises.

Therefore, the aforementioned challenges are aimed to be solved by providing an interoperable and customizable smart contract platform that can be adapted to any business process that could be needed by an Industry 4.0 enterprise or group. The proposed platform is separated from any internal architecture that an enterprise could have, implements private channels and has the ability to access any external data in a secure manner. Thus, a holistic DLT architecture is intended to be achieved, where data integrity and traceability are guaranteed throughout the whole process: from when the data is generated, processed and homogenized, up until it is exploited for business purposes. Finally, the proposed platform is evaluated by implementing an Industry 4.0 use case.

Specifically, the following contributions are provided:

- Design of an interoperable, customizable and fee-less smart contract platform. This contribution addresses the challenges of interoperability, scalability, and cost. Different organizations, using different smart contract platforms, often find it hard to cooperate due to their varying data structures. The proposed design allows for adaptable interfaces to accommodate various business processes, facilitating collaboration among diverse enterprises. Scalability issues are addressed by the customizability of the presented platform, enabling it to handle high-transaction situations without impairing performance. Moreover, the challenge of cost is targeted by ensuring the platform does not require transaction fees, which can often present a substantial burden to businesses.
- Private channel implementation for data privacy. This contribution directly counters the challenge of limited data privacy. As smart contracts execute on distributed ledgers, all data used becomes transparent and accessible, which can

be problematic for sensitive information. By implementing private channels, the proposed platform ensures data privacy for organizations, mitigating this concern.

- Incorporation of oracles for external data access. Addressing the challenge of limited access to external data, the proposed platform incorporates oracles, enabling the smart contracts to access data outside the ledger. This feature expands the operational capability of the smart contracts, ensuring their successful execution.
- Implementation of the proposed platform along with specific use cases. Here, the solutions are translated into practice by developing the platform and demonstrating its functionality in a real-world Industry 4.0 use case developed in collaboration with a leading industrial company (Fagor Automation). This allows a better understanding on how the platform can meet the challenges and needs of real Industry 4.0 enterprises.
- Evaluation of the platform in terms of performance and security. Lastly, this study examines the platform's effectiveness in overcoming the identified challenges.

4.6.2 Motivating Scenario

A consortium of industrial manufacturing enterprises, encompassing raw material providers, component distributors, and finished product manufacturers, faces the escalating necessity to enhance effectiveness, safety, and clarity throughout their supply chain procedures. The intricate structure of contemporary supply chains, coupled with the growing requirement for instantaneous data and comprehensive visibility, has spurred these organizations to pursue inventive approaches to tackle these obstacles. Consequently, these companies acknowledge the transformative power of blockchain and smart contracts in refining their supply chain practices and fostering confidence among all parties involved, such as suppliers, clients, and regulatory bodies.

The motivating scenario of this work revolves around a consortium of these industrial manufacturing companies seeking solutions to address the following challenges:

- Supply Chain Transparency and Traceability: The consortium members aim to track the movement of goods and materials throughout their supply chains. This includes aspects such as monitoring the origin of raw materials, ensuring ethical and sustainable sourcing, and documenting the transportation of goods from suppliers to manufacturing plants, warehouses, and finally to customers.
- Data Integrity and Security: The companies seek a solution that guarantees data integrity, immutability, and traceability, ensuring that no unauthorized changes

can be made to records. This would foster trust among the consortium members and help prevent fraudulent activities, such as counterfeiting or unauthorized component substitutions.

- Improved Collaboration and Partnerships: The consortium members are interested in accessing shared data and analytics that enable them to collaborate more effectively, make better-informed decisions, and optimize their processes. Furthermore, they seek the execution of automated agreements such as pay-per-use schemes and other forms of automated agreements with other actors such as the utilities suppliers (e.g., electricity).

- Regulatory Compliance: The consortium needs a solution that can enforce regulatory requirements automatically, streamlining the compliance process and minimizing the risk of human error or misinterpretation.

- Interoperability and Scalability: The industrial manufacturing companies seek a solution that can be adapted to any business process they require and can seamlessly integrate with their existing systems. This approach ensures that the solution remains flexible and scalable as the companies evolve.

- Data Privacy and Security: The consortium is concerned about protecting sensitive information while still making it accessible to authorized members. They need a solution that addresses data privacy without compromising the benefits of a shared platform.

4.6.3 Business Blockchain for Industry 4.0

The proposed business blockchain architecture is designed for multiple industrial enterprises to interact with external entities in a secure, fast, cheap and straightforward manner. Specifically, the design of the architecture addresses the aforementioned challenges regarding:

- Compatibility or interoperability. The smart contracts and their associated data has to be compatible.

- Performance and scalability. The system must be able to support a relatively high number of simultaneous operations.

- Data privacy. Data privacy must be guaranteed within the blockchain, since business agreements are typically private.

- Access to external data. In order to perform complex business agreements, smart contracts must be able to access external data.
- Incurred costs. A blockchain-based system for business processes should not lead to a considerable increase in costs for companies.

Therefore, in order to tackle these challenges, six main design characteristics are established. Some of the characteristics that are defined such as the DLT type, permissions level and consensus are required in any DLT design, as stated in the work presented by M. Tabatabaei et al. [258]. The aforementioned work provides a comprehensive survey of DLT architectures, advancing in the standardization of DLT design. Apart from the three characteristics that are mentioned above, three more that are not typically mandatory are included; however, they are taken into account in order to solve all the challenges that were mentioned in the introduction.

1. The type of the DLT. An analysis of the existing DLT types has been carried out, while electing the most appropriate type for the given requirements. There are many type of DLTs that have their own particular advantages and drawbacks.
2. The DLT access and permission level (public, private, permissioned, permissionless). Whether the ledger will be public or private and permissioned or permissionless must be established. This design choice impacts several aspects of the DLT such as the level of privacy and its overall performance (i.e., private blockchains are typically much faster than public blockchains).
3. External data access mechanisms. Smart contracts that are unable to access external data are strictly limited, especially when it comes to business agreements. For example, many business agreements require access to external price charts. Therefore, an approach that enable smart contracts to use external data is designed.
4. The interoperability approach. As stated above, the smart contracts and their data must be compatible. Thus, an interoperability approach is designed in order to cover a broad range of applications.
5. Data privacy mechanisms. Data privacy mechanisms within the blockchain are designed, in order to enable private agreements between enterprises.
6. The consensus mechanism. The consensus algorithm is the core of the DLT, since it prevents malicious behavior in the network. Moreover, consensus has a great impact in the DLT performance, thus it has to be carefully designed.

Figure 4.10 shows the overall diagram of the proposed architecture. Inside the white frame we have the starting context, in which several industrial plants share a common blockchain, including homogenized IIoT data. Within the green frame we have the proposed platform, in which many types of enterprises connect with each other and execute automatic agreements based on the blockchain smart contract technology. Industrial plants are connected to this platform via their shared homogenized data blockchain since their data is useful for the execution of the agreements, such as, for example, a pay-per-use agreement for industrial machines with external providers. However, each enterprise has total control over the amount of data that it shares for the execution of smart contract agreements and processes.

A smart contract blockchain platform greatly improves business processes between industrial plants and external suppliers, energy (and other utilities) providers, other production centers and business headquarters by providing a secure, transparent, and decentralized way to manage transactions and interactions between different parties. Smart contracts streamline numerous manual processes currently needed to oversee interactions, including payment tracking, transaction verification, and agreement enforcement. Furthermore, the decentralized nature of blockchain minimizes the risk of fraud and guarantees real-time access to identical information for all involved parties.

DLT Type

In this subsection, the type of DLT that should be implemented in the proposed business-oriented architecture for Industry 4.0 is discussed. As mentioned above, there are some alternative solutions to blockchain, the most popular and widely adopted alternative being DAG based DLTs. However, in this architecture, a smart contract capable blockchain is adopted over the other available options.

One of the main advantages of a blockchain over other types of DLTs, such as DAGs, is its ability to provide a higher level of security, immutability and decentralization. This is achieved through the use of consensus algorithms and cryptographic techniques that make it extremely difficult for any one party to alter or tamper with the record of transactions. Other DLTs have not proven yet to be as decentralized or as secure as blockchains. For example, the IOTA DAG platform currently has a centralized coordinator node, and it has experienced many security issues over the last few years [260].

Another advantage of a blockchain is its ability to facilitate smart contract functionality, which is crucial in this work. Smart contracts are stored and replicated on the blockchain network. This allows for the automation of certain processes and the ability to easily track and verify the execution of the contract. In an industrial context, this

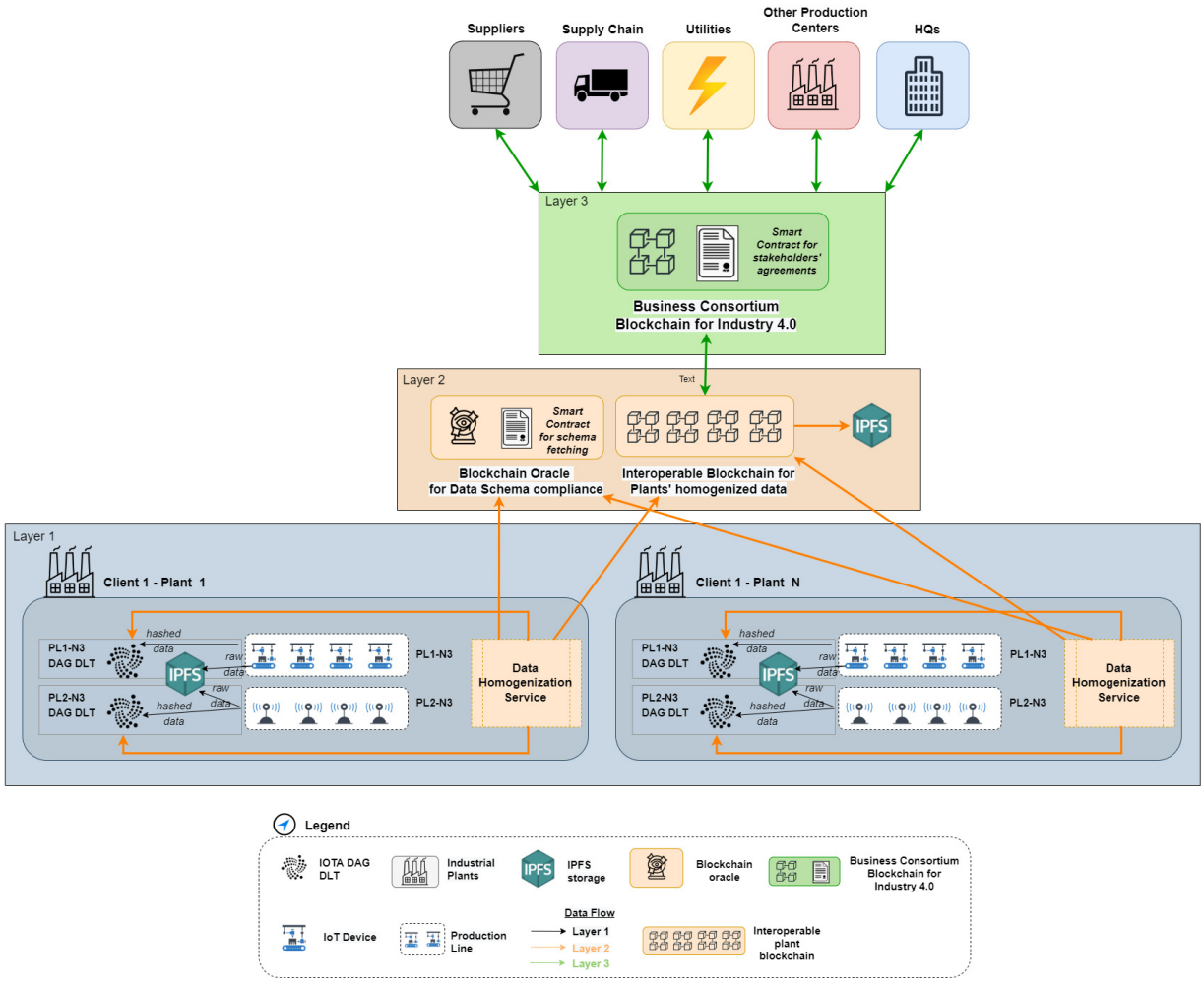

Fig. 4.10 Overall diagram of the proposed platform. In white, the starting context [259]. In green, the proposed platform. Arrows represent the data flow.

enables the efficient and secure execution of transactions between different parties in the business ecosystem.

Finally, even though DAGs are known to be much faster than the most popular blockchains, such as Bitcoin, there are some business-oriented permissioned blockchain solutions that also offer a high transaction processing speed. For example, the Hyperledger Fabric blockchain can achieve hundreds or even up to 1000 transactions per second. Nonetheless, in business environments where digital contracts are executed, the required throughput is not as high as it might need to be at lower levels where the IIoT devices are located [23].

Network accessibility and permissions

In this subsection, whether a public-permissionless or private-permissioned (consortium) blockchain is the most appropriate choice for the proposed scheme is studied.

While public blockchains have many advantages, such as being totally decentralized and providing transparency, they may not be the best choice for the proposed architecture of industrial plants and other entities collaborating and conducting business transactions. Overall, the drawbacks of public blockchain can be summarized as follows [261]:

- Lack of privacy. Public blockchains are open to anyone, which means that sensitive business information may be exposed to competitors or malicious actors.
- Slow transaction processing. Public blockchains have inefficient consensus algorithms, which can lead to longer transaction processing times.
- Lack of scalability. With a large number of participants, public blockchains can become congested, leading to a lack of scalability.
- Lack of governance. With anyone participating, it can be difficult to make decisions about the network and manage it effectively.

On the other hand, consortium blockchains are a more suitable choice in this type of business environment due to the following reasons [261]:

- Increased security. By limiting the number of participants, the risk of malicious actors infiltrating the network is reduced. This ensures that only trusted entities can access and validate transactions.
- Faster transaction processing. With a smaller number of participants, the consensus mechanism can be faster, leading to faster transaction processing times.
- Better scalability. As the number of participants is limited to the necessary parties only, the network can handle a larger volume of transactions, making it more scalable.
- Better privacy. By allowing only specific participants to access the network, sensitive business information can be kept private.
- Better governance. With a smaller group of participants, it is easier to manage and make decisions about the network.

- Smart contracts capabilities. Most consortium blockchains also provide the ability to execute smart contracts, which can be used to automate business processes and streamline workflows.

- Customization. Consortium blockchains can be customized so that they comply with the specific requirements and necessities of the businesses that are involved.

Overall, a public blockchain is not appropriate for this type of environment because it lacks privacy, has slow transaction processing, lacks scalability and lacks proper governance. A consortium blockchain, on the other hand, is a better option as it allows for a more controlled, private, and efficient network with the ability to provide custom features that are needed in this scenario.

External Data Access

As stated before, one of the significant challenges that smart contracts are facing in blockchain systems is accessing trustworthy external data. Thus, in this work blockchain oracles are used in order to enable smart contracts access external data and improve the usability of the architecture, since oracles provide more reliability, automation, transparency, and interoperability. The oracle mechanism involves the identification and establishment of trusted data sources, the appointment of a set of nodes as oracles, and the use of smart contracts to enforce conditions based on the data provided by the oracles. By having a trusted set of oracle nodes retrieve data from trustworthy sources, the accuracy of the information used in the smart contracts is assured. This, in turn, allows for the automation of contract execution and enforcement, adding transparency to the process as all data can be audited. Figure 4.11 shows a graphical representation of blockchain oracles and smart contract powered blockchains.

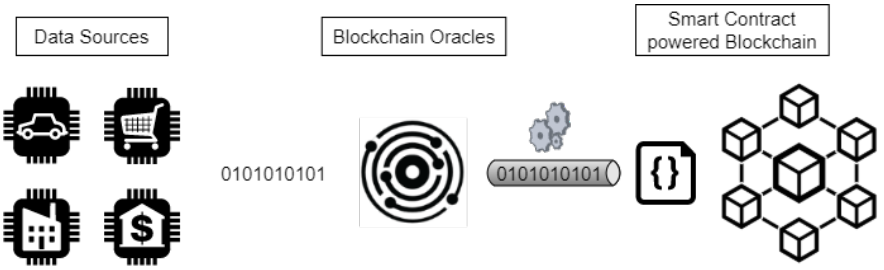

Fig. 4.11 Blockchain oracles retrieving external data for smart contacts.

In the proposed consortium Industry 4.0 environment, oracles are intended to be used to retrieve relevant information such as market prices, product availability, weather

conditions, etc. For example, in an Industry 4.0 supply chain, the oracles can retrieve trustworthy external information on the current price of electricity. Furthermore, in a manufacturing setting, oracles can also retrieve information on the availability of raw materials, triggering the smart contract to order more materials or change production schedules based on the availability. In this way, oracles provide real-time data that can be used to automatically enforce the terms of smart contracts and expand their usability.

Interoperability

In this subsection, an interoperability connector gateway is designed to enable communication between the business blockchain and other blockchains that contain the data from each business. An interoperability gateway connector for the proposed blockchain platform is a software component that allows different blockchain networks to communicate and share information with each other. Connectors enable transferring of assets or data between networks and execution of smart contracts across multiple networks. Thus, the connector acts as a bridge between the different networks, allowing them to interoperate and work together seamlessly. In the case of industrial plants, the proposed business blockchain has to be connected to the blockchain that contains homogenized plant data, as specified in [259]. In conclusion, this interoperability solution is crucial since smart contracts need data in order to function correctly.

The presented design for the interoperability gateway is comprehensive and multifaceted. Figure 4.12 depicts the architecture of the interoperability gateway, along with its components. The gateway is divided into three layers (L1, L2 y L3). The bottom layer (L3) include the core components and provide the basic functionalities of the gateway: the rules engine, the mapping engine and the communication layer. On top (L2) we have the adapter itself, which connects to the core parts from L3 and enables the actual communication between the DLTs. Finally, on the top (L1) we have the monitoring system, which acts as the user interface of the whole gateway.

The gateway begins with the management and monitoring component that oversees the connector, managing its configuration, logging, and continually tracking the health and performance of the adapter.

The adapter is the next vital piece of the puzzle, connecting various blockchain networks and acting as a translator for data and commands between them. It is adaptable and supports multiple blockchain protocols based on architectural requirements.

The rules engine follows, enforcing any applicable business logic or rules during data exchange across different blockchain networks. It's a crucial tool for ensuring compliance rules and implementing smart contract functionality.

Then we have the mapping engine, which takes on the role of a data and assets medi-

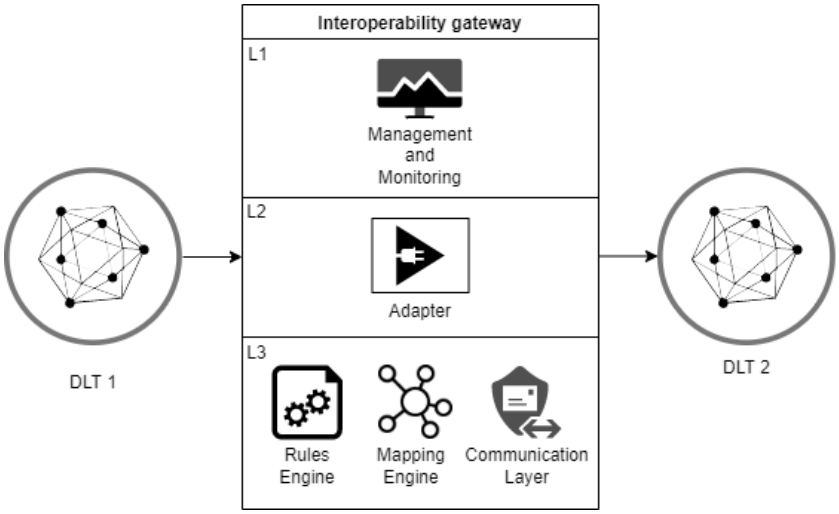

Fig. 4.12 Interoperability gateway scheme

ator among different networks. It manages intricate tasks like converting data formats, mapping smart contract functions, and handling any data validation or transformations necessary.

The communication layer is another essential part of the design, responsible for secure inter-network communications. It takes care of encrypting and decrypting data, establishing secure connections, and tackling network failures.

Lastly, an API for the connector is provided, making it accessible and communicative to external applications. This final touch ensures that the interoperability gateway is not just secure and robust, but also user-friendly and efficient.

Data privacy

Data privacy is a major concern in consortium blockchains as they involve multiple parties sharing sensitive information on a common platform. Private channels help address this issue by allowing communication between a select group of participants [262]. This feature enables organizations to securely and privately exchange sensitive information without revealing it to the entire network. Specifically, data shared within a private channel is encrypted and solely accessible to channel members, adding another layer of security. This approach contributes to the preservation of sensitive information confidentiality and fosters trust among organizations. Private channels utilize a mix of symmetric and asymmetric encryption, including Advanced Encryption Standard (AES) for data encryption, ECDSA for digital signatures, and Transport Layer Security

(TLS) for secure network communication.

An alternative privacy method in consortium blockchains is homomorphic encryption [263]. This encryption technique allows for computations on encrypted data without the need for prior decryption. It offers privacy protection by enabling the processing of sensitive information while remaining encrypted, preventing unauthorized access. However, homomorphic encryption is still in its early development stages and faces several challenges, such as high computational overhead and a narrow scope of practical use cases.

In the case of the proposed Industry 4.0 consortium blockchain, private channels offer a simpler and more practical solution for privacy protection compared to homomorphic encryption. Private channels allow for secure communication between a select group of participants, which meet the privacy requirements for most use cases in industry 4.0. Additionally, private channels are easier to implement and require less computational overhead compared to homomorphic encryption, making them a more feasible option.

Figure 4.13 depicts the private channels architecture. Peer nodes 1 and 3 belong to one channel (blue) whilst peer node 2 belong to other channel (red). Each channel includes its own separate ledger and smart contract. The orderer node keeps the order of proposed transactions, validates endorsement signatures, and broadcasts messages to peers.

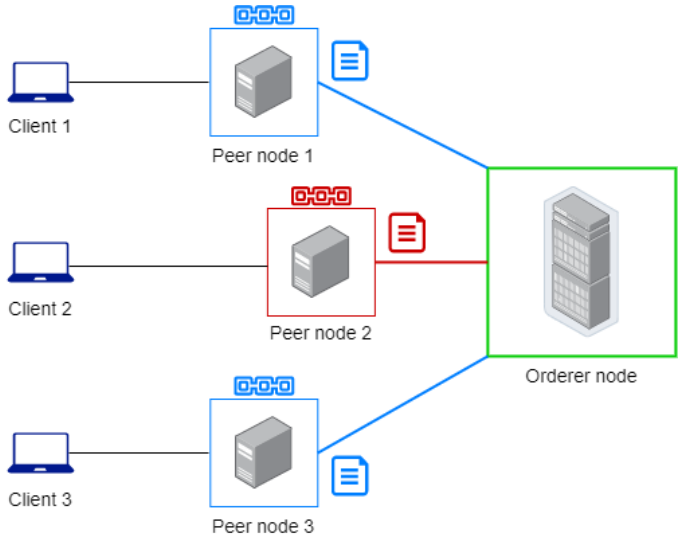

Fig. 4.13 Private channels architecture. The red and blue channels have their own separate ledger and smart contracts.

Consensus

When it comes to choosing a consensus algorithm for a consortium blockchain, there are many factors to consider, such as performance, security, scalability, and reliability. In this work, the use of Raft is proposed.

Raft is a leader-based consensus algorithm that is widely used in distributed systems. It provides a high level of fault tolerance, which means that even if some nodes in the network fail, the blockchain can continue to operate normally. Additionally, Raft has proven to be highly scalable, making it suitable for use in large, complex networks.

Figure 4.14 depicts the phases of the Raft consensus. When a system starts or the leader fails, a new leader is elected. The leader accepts client requests, appends new entries to its log, and replicates the log entries to follower nodes. Once a log entry is replicated to the majority of nodes, it is marked as "committed", and the nodes execute the operation. Furthermore, to maintain storage efficiency, the algorithm employs log compaction and snapshots, which involve compressing the current log and discarding old log entries.

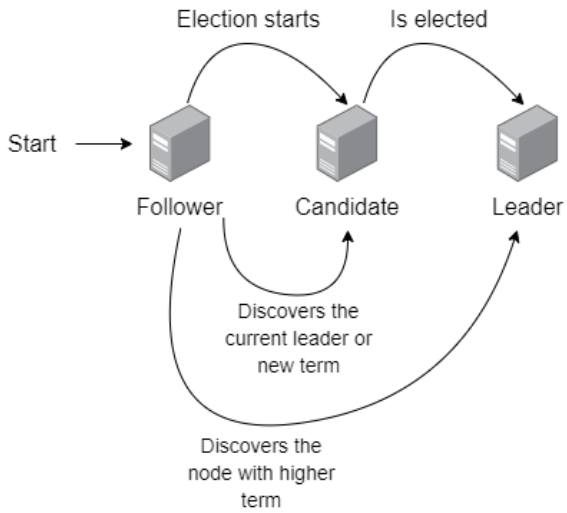

Fig. 4.14 Raft consensus mechanism phases

One of the key advantages of Raft over other consensus algorithms is its simplicity. Unlike other consensus algorithms, such as PBFT, which can be difficult to implement and understand, Raft is straightforward and well-documented. This makes it easier to build and maintain a Raft based blockchain.

Furthermore, Raft is generally considered to have better performance than PBFT, particularly in scenarios with larger networks. PBFT requires more message exchanges

and processing time due to its complex message validation process, which becomes a bottleneck in large networks.

Another important factor in choosing a consensus algorithm is security. Raft provides strong security guarantees. In particular, Raft's leader-based approach makes it difficult for attackers to disrupt the consensus process, helping to prevent attacks such as the "double spend" issue. However, in comparison with PBFT, Raft is better suited for scenarios where the primary concern is handling node failures and network partitions rather than malicious behavior. Nonetheless, in the proposed network the presence of a large number of malicious nodes is unlikely since it is a permissioned consortium network between enterprises.

In conclusion, with its combination of reliability, scalability, and security, Raft provides a strong foundation for secure and efficient data sharing among the various stakeholders in the network.

4.7 Implementation and Evaluation

This section details the implementation and evaluation of the proposed architecture. The section is organized into three distinct parts. First, the implementation and evaluation of the enhancements to Layer 1, the "Data Source Layer" are introduced. Following this, the implementation of the "Bridge Layer" is discussed, where a process for industrial data homogenization is proposed, facilitating its subsequent storage and processing within an interoperable plant blockchain. The Bridge Layer also incorporates a monitoring system to ensure the architecture operates as intended. Lastly, the implementation of the entire architecture is described, encompassing the previously mentioned layers along with the third layer, the "Business Layer". This final section involves connecting multiple entities within an Industry 4.0 consortium. As part of the implementation, two use cases are showcased: the application of the architecture to measure the effectiveness of industrial machines (OEE), and a realistic implementation developed in collaboration with a leading industrial company, Fagor Automation.

4.7.1 Data Source Layer

Security Analysis

In this section, the security implications of the proposed improvements to the DAG DLT are analyzed, focusing on IPFS storage, cryptographic algorithms, and the anti-spam reputation mechanism.

Regarding IPFS storage, the use of IPFS ensures that if a malicious node manipulates the stored data, the corresponding hashes will change, allowing for the quick detection of the attack. Since the hashes and metadata of the files are stored in the theoretically immutable DAG DLT [211], the data cannot be altered. Consequently, manipulating the IPFS hashes would be infeasible.

In terms of cryptographic algorithms, although a comprehensive security evaluation of Quark and FALCON is beyond the scope of this work, the literature provides evidence of their security properties. In the analysis performed by Aumasson et al. [235], Quark is shown to be secure against various types of attacks, with the lightest instance (u-Quark) conjecturally providing at least 64-bit security. Additionally, FALCON [240] uses a true Gaussian sampler, ensuring negligible leakage of information on the secret key, and meets the highest NIST security level (level 5).

Lastly, the security of the proposed reputation-based anti-spam mechanism in two scenarios is examined: a compromised honest device within the DAG network, and a new malicious device joining the network. The second scenario is less likely in an industrial context due to stringent physical security measures.

If an honest device is compromised, consider a device D_2 with a history of T = 50 and V = 48, resulting in a reputation score R = 9.6 $\approx$ 10. If D_2 is compromised and emits 100 invalid transactions, T and V will change to 150 and 48, respectively, causing R to drop to 3.2 $\approx$ 3. This forces the malicious device to perform substantially more PoW before sending data. As R continues to decrease, the compromised device will eventually be unable to participate in the network.

If a new malicious device joins the network, a device D_3 with T = 0 and V = 0 will initially have a reputation score R = 5. If the malicious device D_3 emits 50 invalid transactions, T will increase to 50 while V remains at 0. This results in R dropping from 5 to 0, effectively removing the malicious device from the network.

In conclusion, the proposed improvements to the DAG DLT ensure a secure and resilient architecture capable of detecting and mitigating potential security threats in IIoT environments.

Performance Evaluation

In this section the performance evaluation of the proposed DAG DLT for the data source layer of the architecture is presented. The purpose of the experiments is to validate the proposed improvements over the state-of-the-art transaction-based DAGs (e.g. IOTA). A methodology that is commonly used in the field of computer science [264] is employed. The experiments have been performed on the following aspects:

- Public key algorithms signature and verification delays. Its impact on the DAG

throughput has been also studied.

- Hash functions delays and their impact on throughput.
- The impact of the reputation-based anti-spam mechanism. The time nodes had to perform PoW before issuing transactions has been measured.
- IPFS storage reduction and impact on throughput.
- The impact of SSD swap memory in the performance.

The experiments have been performed using the Python-based IOTA DAG simulator "DAGsim" [265]. The aforementioned simulator was installed on a laptop with an Ubuntu operating system, an i7 processor, 16 GB of RAM and a 512 GB SSD drive. In order to carry out the experiments, the simulator has suffered several modifications in order to correctly represent the baseline. These modifications are as follows:

1. The simulator has been set to use the cryptographic functions that are used in IOTA's most recent version "Chrysalis": EdDSA and Blake2b.
2. With the new "Chrysalis" version of IOTA, a new transaction can now reference up to eight previous transactions, instead of only two.
3. The concept of "parent" that was introduced in the Chrysalis update has been added. Now a transaction must reference its "parents" (i.e. its predecessors).
4. A random data generator has been implemented.
5. The PoW protocol has been implemented. Before issuing a transaction, a node has to perform some PoW.

Finally, the proposed improvements from Section 4.3.2 on top of the baseline have also been implemented: the cryptographic functions (Quark and FALCON), the IPFS storage solution, the reputation mechanism for PoW and the SSD swap memory boost for IIoT devices.

The following simulator configuration has been defined:

- Transactions: 100
- Rate of transactions (λ): 10
- Number of nodes: 4
- Distance between nodes: 1

- Tip selection algorithm: weighted random walk using the Markov Chain Monte-carlo (MCMC) algorithm

Each experiment has been conducted three times and calculated the average value of the obtained values in order to reduce the influence on the results of other processes within the operating system.

Hash Function Analysis: The generation delays of various hash functions and their subsequent effect on the throughput of the DAG have been examined. As can be seen in Figure 4.15 (a), the proposed Quark algorithm significantly outperforms Blake2b in terms of delay, exhibiting a 152% improvement. This reduced delay can be attributed to Quark's more efficient computation and lighter-weight design.

Furthermore, Figure 4.15 (b) demonstrates the considerable impact of Quark on throughput, as it increases by 50% when compared to Blake2b, rising from 66 transactions per second (tps) to 99 tps. This enhancement in throughput can be traced back to the PoW algorithm employed within the DAG, which is optimized for use with Quark.

The superior performance of Quark in both delay and throughput can be attributed to its design, which focuses on streamlining computations and minimizing resource consumption. This efficiency translates into a more scalable and effective system for handling transactions within the DAG. In summary, the Quark algorithm's lower generation delays and higher throughput make it a promising option for improving the performance and scalability of DAG-based systems.

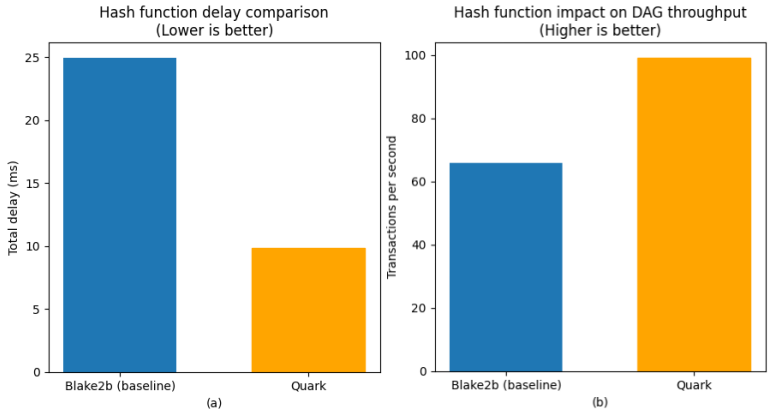

Fig. 4.15 Hash performance comparison. (a) Hash delay (b) Hash impact on throughput

Digital Signature Algorithm Analysis: The signature and verification delays of different digital signature algorithms have been compared. As depicted in Figure 4.16 (a), EdDSA demonstrates a considerably faster signing process, being 200% faster than

FALCON. This can be attributed to the more efficient and streamlined design of EdDSA in generating signatures.

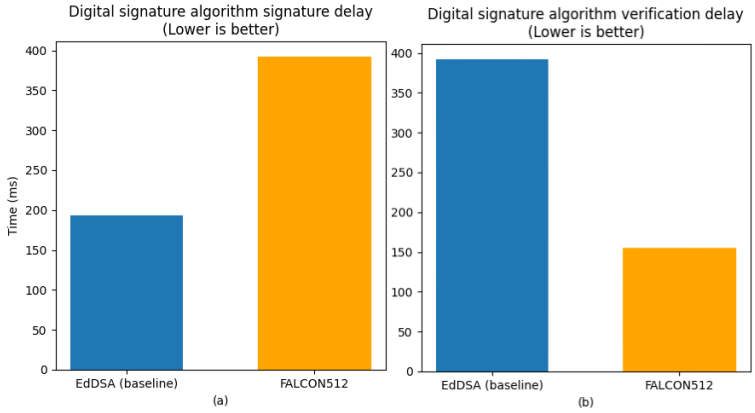

Fig. 4.16 Digital signature algorithm (a) Signature delay (b) Verification delay

On the other hand, FALCON outshines EdDSA in the verification process, as it exhibits a 152% lower delay. This performance advantage can be credited to FALCON's advanced verification algorithm, which is optimized for rapid and accurate validation of signatures. Despite the differences in signing and verification delays, EdDSA is only slightly faster than FALCON overall, as anticipated. It is essential to note that while FALCON may have a minor performance disadvantage compared to EdDSA, it compensates for this by being quantum-resistant, providing a higher level of security against potential quantum computing attacks.

In conclusion, while EdDSA offers a slightly faster overall performance, FALCON's quantum-resistant capabilities and minimal impact on throughput make it an attractive alternative for securing systems against potential future threats posed by quantum computing advancements. The growing concern about quantum computing's potential to break current cryptographic schemes has made the development and adoption of quantum-resistant algorithms crucial for ensuring long-term security.

Lastly, as illustrated in Figure 4.17, FALCON has a negligible negative impact on throughput, at just 3%. This minimal effect on throughput demonstrates that FALCON remains a viable option for systems that require quantum-resistant security without significantly sacrificing overall performance. The ability to maintain a high level of throughput is particularly important in real-world applications where transaction processing and data transfer speed are crucial factors, such as IoT devices.

Storage: The experimental results, as depicted in Figure 4.18, demonstrate that utilizing the IPFS for data storage can significantly alleviate the storage burden associ-

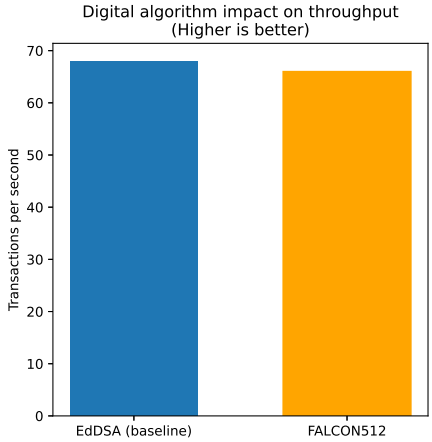

Fig. 4.17 Digital signature algorithm impact on throughput

ated with DLTs. By comparing the baseline DAG with this improved version, we can observe substantial differences in storage requirements as the number of transactions increases. At 100 transactions, the baseline DAG occupies 14.33 MB of storage, while this improved version, which leverages IPFS, requires a mere 34.59 KB. This represents a remarkable 99.76% reduction in storage size, illustrating the effectiveness of the proposed solution. As the number of transactions grows to 1,000, the size of the baseline DAG expands to 143.55 MB. In contrast, the presented improved version that employs IPFS only increases to 356 KB. This still corresponds to a significant 99.75% reduction in storage size when compared to the baseline DAG.

Before, the possible impact of IPFS in the performance of the DAG due to its querying ineffectiveness has been discussed. According to the results, the baseline storage version achieved an average of 66 tps, while the IPFS version achieved an average of 47. However, the performance issues of IPFS have already been studied in [266], where the authors propose a promising solution on this aspect. Improving IPFS performance is out of the scope of this work, thus, this aspect has to be worked on in the future. In conclusion, in this work the aim is to encourage as much as possible the participation of lightweight devices in the DAG DLT. Therefore, in this case, the enormous reduction of storage burden for IIoT devices offsets the loss of some throughput capacity.

Anti-spam reputation mechanism: The amount of PoW that the nodes perform for each transaction has been tested and the total simulation time was registered. For each experiment the same amount of reputation for all nodes was established, so that the results could be comparable. First, the DAG DLT with no reputation system (i.e.

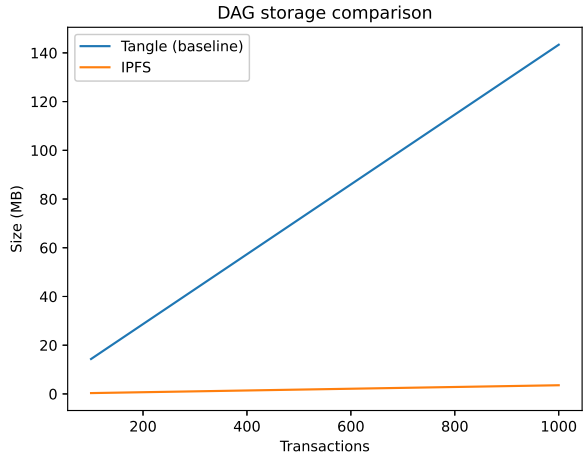

Fig. 4.18 Storage comparison

baseline) have been tested, and then the improved version with high and maximum reputation scores have also been tested. For the baseline version the amount of PoW a device has to perform was set to be the same as on a device with a medium reputation of five. As shown in Figure 4.19, higher reputations greatly reduce the time a device has to perform PoW, thus they greatly reduce the computational burden and the energy consumption of the DLT. The difference between medium / no reputation and high reputations is determined by the exponential variation of the computational effort to be made when the PoW difficulty increases.

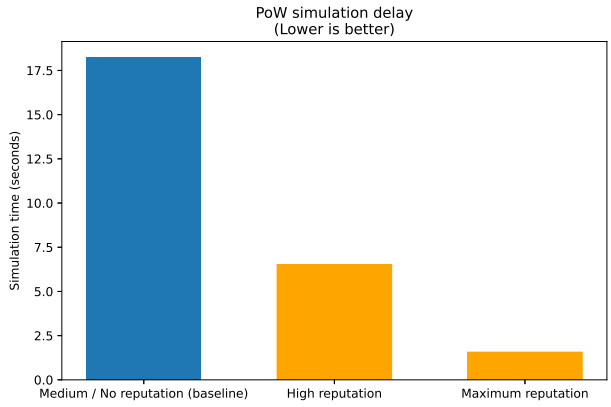

Fig. 4.19 Anti-spam reputation mechanism

4.7.2 Bridge Layer

Data Homogenization Process with Decentralized Oracles Implementation

The machine data that is employed in this prototype is based on a real-world JavaScript Object Notation (JSON) structure that was obtained from actual industrial sensors. The IIoT devices from the simulated scenario collect data on the performance of the production line, the quality of the products being produced, timestamp data, diagnostics, and many other factors. These data can be used to optimize the production process and improve efficiency. When implementing the prototype, heterogeneous data similar to a real-world environment that was described in Section 4.4.1 is simulated. Thus, this implementation aims to solve the challenges related to the security, integrity, and heterogeneity of industrial data.

Specifically, the following Industry 4.0 IIoT equipment is simulated:

- Smart sensors: These sensors can collect and transmit data about the performance and operation of machines, processes, and systems in real time.
- Predictive maintenance systems: These systems use machine learning and data analytics to predict when maintenance is needed, helping to reduce downtime and improve efficiency.
- Robotic systems: These systems can automate tasks such as material handling, assembly, and inspection, helping to increase productivity.

IOTA is used as the production line DAGs to process the raw data since IOTA is currently known to be the most advanced DAG DLT solution [267], especially in terms of performance. As for the oracle service, there are many relevant options that can be chosen. As mentioned before, the most well-known oracle platform is ChainLink[4], which is focused on deploying Ethereum-compatible oracles.

However, in this work, the Ethereum blockchain is not used since it lacks interoperability capabilities, along with low-performance capabilities. Furthermore, to provide interoperability, Polkadot[5] has been chosen as the oracle service, as well as the blockchain solution in which the homogenized data will be stored. In this case, a relay chain in which the homogenized data is stored has been implemented, along with a parachain that acts as an oracle service.

This implementation leaves the possibility of extending the functionality of the architecture by connecting other parachains in the future, which for example, could carry

[4] `https://chain.link/`
[5] `https://www.polkadot.network/`

out the execution of smart contracts that could establish business relationships with other entities (i.e., suppliers, customers, etc.).

Finally, the JSON-based Eclipse Unide data model is being used, as shown in Listing 1. The Unide data model is specifically designed for manufacturing processes, and it is trusted by several major parties, such as SAP or Bosch.

Listing 1 Eclipse Unide data model

```
{
  ''type'': ''object'',
  ''properties'': {
    ''content-spec'': {
      ''type'': ''string'',
      ''default'': ''urn:spec://eclipse.org/unide/machine-message#v3'',
      ''description'': ''Defines what the format version is''
    },
    ''device'': {
      ''$ref'': ''definitions.json#/definitions/device''
    },
    ''part'': {
      ''$ref'': ''definitions.json#/definitions/part''
    },
    ''measurements'': {
      ''allOf'': [
        {
          ''$ref'': ''definitions.json#/definitions/measurements''
        },
        {
          ''items'': {
            ''properties'': {
              ''series'': {
                ''required'': [
                  ''time''
                ]
  },
  ''required'': [
    ''content-spec'',
    ''device'',
    ''measurements''
  ]
}
```

First, a NodeJS client that emulates several industrial devices and periodically sends industrial raw data to an IPFS file system has been implemented. Then the resulting IPFS hash is sent to the IOTA DAG DLT. Afterward, the data homogenization client in NodeJS has been implemented. This client performs the following sequence of six tasks:

1. Access the IPFS raw data using the hash that is stored in the production line IOTA DAG DLT. An example of an industrial raw data JSON is shown in Listing 2.

2. Request the oracle service to retrieve the data model. Figure 4.20 shows the retrieval of the data model by the Polkadot parachain that was set as oracle.

3. Perform the data homogenization process. The mapping between the raw data schema to the standard Eclipse Unide data model schema was made using the *jsonpath-object-transform* NPM package. Listing 3 shows the NodeJS code of the transformation process of the data according to the Unide model.

4. To assure that the process was correctly executed, the resulting JSON was validated using the Ajv JSON schema validator.

5. Add the used data model and the resulting homogenized data JSON to IPFS. An example of the homogenized raw data is shown in Listing 4.

6. Send a transaction to the Polkadot relay chain (interoperable plant blockchain) to store the IPFS hash of the homogenized data. Figure 4.21 shows the stores IPFS hash pointer of the homogenized data within the Polkadot blockchain.

Listing 2 Raw industrial data JSON example

```
{
  ''device'': ''20131''
  ''metadata'': { ''origin'': ''StrokeData'' },
  ''keys'': {
    ''id_stroke'': 4705340,
    ''id_die'': 18,
    ''id_die_string'': ''69-14'',
    ''dipartcounter'': 4704419
  },
  ''data'': [
    {
        ''filter'': true,
        ''cs_workmode'': 5,
        ''cs_partcntr_shift1'': 2,
        ''cs_oee'': 95,
        ''ts'': ''2019-07-04T13:33:03.969Z'',
        ''series'': [Object]
    }
  ]
}
```

Listing 3 Data transformation in NodeJS code

```
const schema = dataModel;

var transform = require('jsonpath-object-transform');
var template = {
    ''type'': '''',
    ''content-spec'': '$.metadata.origin',
    ''device'': {
        ''id'': '$.device'
    },
    ''part'': '$.keys',
    ''measurements'': '$..data'
}
const homogenizedData = transform(IPFSRawdata, dataModel, template);
```

Listing 4 Homogenized industrial data JSON according to the Eclipse Unide model

```
{
  ''type'': ''object'',
  ''content-spec'': ''StrokeData'',
  ''device'': {''id'': ''20131''},
  ''part'': {
    ''id_stroke'': 4705340,
    ''id_die'': 18,
    ''id_die_string'': ''69-14'',
    ''press_vel'': 17.1,
    ''isstrokecalssification'': 2,
    ''bvalidstroke'': false,
    ''dipartcounter'': 4704419,
    ''id'': 98
  },
  ''measurements'': [
    {
        ''filter'': true,
        ''cs_workmode'': 5,
        ''cs_partcntr_shift1'': 2,
        ''cs_partcntr_shift2'': 0,
        ''cs_partcntr_shift3'': 0,
        ''cs_availablesamples'': 1180315,
        ''cs_productionsamples'': 1110909,
        ''cs_measuredsamples'': 4401216,
        ''cs_oee'': 95,
        ''ts'': ''2019-07-04T13:33:03.969Z'',
        ''series'': [Object]
    }
  ]
}
```

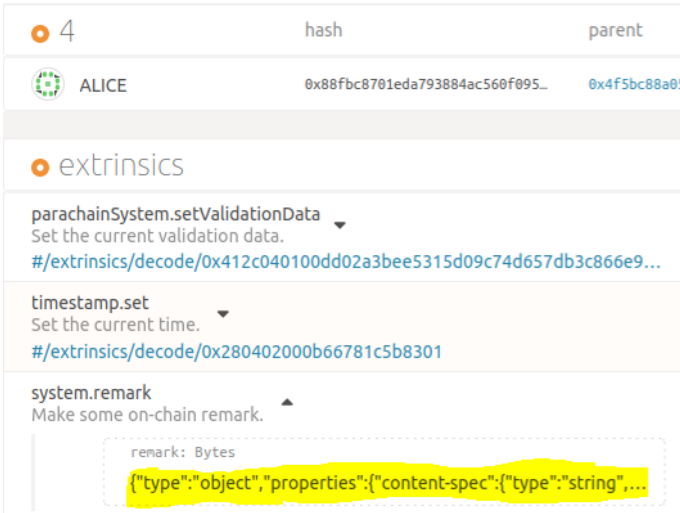

Fig. 4.20 The data model (highlighted) after being retrieved by the Polkadot parachain oracles.

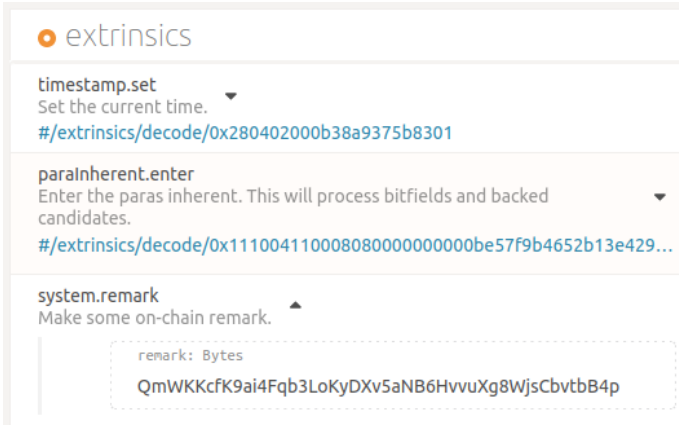

Fig. 4.21 The reference of the homogenized data (system.remark) within the Polkadot relay chain.

Monitoring System Implementation

In this subsection, the implementation of the monitoring system that was designed for the proposed architecture is presented. In this implementation, NodeJS and ExpressJS are used for data retrieval to provide compatibility with the rest of the architecture. For this preliminary version, an API that includes information on the four modules that are explained in Section 4.4.2 was designed. To properly show the monitoring data, the

ELK Stack was employed. The aforementioned tools enable advanced real-time data visualization and monitoring with an easy-to-use dashboard. Thus, creating a dashboard from scratch was not necessary, which would have been a highly complex process. The ELK stack has been proven to be an ideal solution for the given needs, as shown in other relevant works [268].

In the implemented API modules, the following information is shown:

1. **The IIoT data monitoring.** This module shows several metrics that are related to the raw data that comes from industrial machines. The monitoring probes are set directly at the sensor level when the data is generated. The total number of devices within the industrial plant, the number of active devices, the percentage of active devices, the number of sent messages (i.e., raw data transactions), the data generation rate, and the average temperature of the devices are measured. Listing 5 shows an example of the returned IIoT metrics from the monitoring API.

2. **The DLTs monitoring.** This module shows several metrics that are related to IOTA (production line DLT) and Polkadot (interoperable plant blockchain). It shows the overall throughput of each DLT, the transaction validation times, the associated costs (if any), information about the peer nodes, the consensus model, throughput, number of blocks, smart contract information (if any), etc. Listing 6 shows a trimmed example of the returned plant blockchain metrics from the monitoring API.

3. **The oracles monitoring.** This module shows several metrics that are related to the oracles. It shows which oracles have been used the most, which are currently available, the throughput capacity, the accumulated usage fees, the latest retrieved data, the quality of the data, etc. The "quality of data" metric shows whether the retrieved data model JSON is valid or not. Listing 7 shows a trimmed example of the returned blockchain oracles metrics from the monitoring API.

4. **The storage monitoring.** This module shows several metrics that are related to the storage of the data within the IPFS file system, such as performance, storage usage, peer nodes information, the generated hashes, version, IP addresses, etc. Listing 8 shows a trimmed example of the returned IPFS storage metrics from the monitoring API.

Listing 5 IIoT devices monitoring API data example

```
{
''Total devices'':41,
''Number of active devices'':27,
''Devices ID list'':[
   [
      77579,
      56457,
      42678,
      90564,
      35677,
      38909,
      38322,
      98532,
      ...
   ]
],
''Percentage of active devices'':66,
''Number of sent messages'':41,
''Data generation rate each second'':2.7,
''Average temperature'':36
}
```

Listing 6 Polkadot plant blockchain monitoring API data example

```
{
   ''Validators'':{
      ''address'':''5GNJqTPyNqANBkUVMN1LPPrxXnFouWXoe2wNSmmEoLctxiZY'',
      ''balance'':''999,997,674,890,367,678'',
      ''nonce'':''478''
   },
   ''Account nonce'':''89'',
   ''Last block timestamp'':''1668064292'',
   ''Chain Info'':{
      ''ss58Format'':34,
      ''tokenDecimals'':[
         12
      ],
   },
   ''Account nonce'':''127'',
   ''Last block timestamp'':''1664197980005'',
   ''Blocks'':''4114'',
   ''Current throughput'':979,
   ''Max throughput capacity'':997,
   ''Smart Contracts'':''No smart contracts found''
}
```

Listing 7 Polkadot oracles monitoring API data example

```
{
   ''Validators'':{
      ''address'':''5GrwvaEF5zXb26Fz9rcQpDWS57CtERHpNehXCPcNoHGKutQY'',
      ''balance'':''999,993,268,520,263,875'',
      ''nonce'':''108''
   },
   ''Account nonce'':''47'',
   ''Last block timestamp'':''1664198880029'',
   ''Chain Info'':{
      ''ss58Format'':42,
      ''tokenDecimals'':[
         12
      ],
   },
   ''Blocks'':''1047'',
   ''Current throughput'':981,
   ''Max throughput capacity'':1003,
   ''Number of active oracles'':4,
   ''Latest retrieved data'': {...},
   ''Accumulated fees'':''0.013 EUR'',
   ''Quality of the data'':''Good''
}
```

Listing 8 IPFS storage monitoring API data example

```
    {
   ''IPFS ID'':[
      {
         ''id'':''12D3KooWRuhwh6FSafpj88cBYZsTzprU1pybo9hPbcNddniPXQbE'',
         ''publicKey'':''CAESIO8ZQSXXfe3JQ3RHxnuBP9BiZjiRCoYzhsckxj81tHHt'',
         ''addresses'':[
            ''/ip4/10.0.2.15/tcp/4001/p2p/12D3KooWRuhwh6FSafpj88cBYZsXQbE''
         ]
      }
   ],
   ''IPFS version'':''0.13.0'',
   ''Config'':[
      {
         ''Datastore'':{
            ''HashOnRead'':false,
            ''StorageGCWatermark'':90,
            ''StorageMax'':''10GB''
         }
      ],
      ''Repo stats'':[
         {
            ''numObjects'':4345,
            ''repoSize'':25088708
         }
      ]
   }
```

Results

In this subsection, the gathered results from the monitoring system based on a test run of the data homogenization architecture over several days is shown. However, after running the process for several days, a 12 to 14h simulation generates sufficiently robust and realistic results. Thus, major variations in longer simulations have not been observed. The simulated smart factory includes a total number of 500 IIoT devices that send random data at a random rate using the IoT-sim package. During the tests, the number of active IIoT devices varies randomly to simulate a realistic scenario.

The main purpose of the simulations is to demonstrate the viability and security, and performance sufficiency of the architecture. Furthermore, the use of the designed monitoring system also demonstrates its usefulness.

The raw data is processed by IOTA and IPFS at the production line level, and then it is homogenized and processed by a Polkadot plant blockchain. A Polkadot parachain network of a random number of active oracles from a total number of ten is also set. The simulation has been executed using a computer with an i7 9th generation CPU, 16 GB of RAM, and an SSD drive. Several Kibana graphs showing the following metrics generated from the monitoring system have been generated:

- **IIoT devices.** The number of active devices (Figure 4.22), the average temperature (Figure 4.23) and the Overall Equipment Effectiveness (OEE) (Figure 4.24). By generating these graphs, several aspects can be analyzed, such as the production flow, identifying possible device failures, overheating problems, and optimize the effectiveness of the industrial equipment by utilizing data-driven techniques as shown in [269].

- **Storage.** The number of raw data JSONs that are inserted in IPFS from the IIoT devices are measured, and compares with the data that is finally processed by the IOTA DLT (i.e., processed JSON hashes in IOTA), as shown in Figure 4.25. These measurements could help us identify possible anomalies regarding the generating of the data from the IIoT devices. There is also the comparison of the size of the data inside IPFS compared to the amount of size of the processed IPFS hashes in IOTA, as shown in Figure 4.26. The data size monitoring could be useful to optimize storage space and also visualize the enormous storage burden that is displaced from the DLT by using decentralized IPFS storage.

- **The DLTs.** The average throughput of IOTA and Polkadot during the simulation, as shown in Figure 4.27. Measuring the throughput of the DLTs is crucial in terms of data flow optimization and bottlenecks avoidance [270].

- **The oracles.** The average number of oracles was measured during the simulation, as shown in Figure 4.28. By analyzing the number of blockchain oracles that are involved in providing external data to the architecture, we are able to determine the degree of centralization of the system. For example, having only one active oracle would imply a high degree of centralization, which could affect the security of the whole industrial architecture. Furthermore, the number of active oracles is also useful when calculating the associated costs of this service.

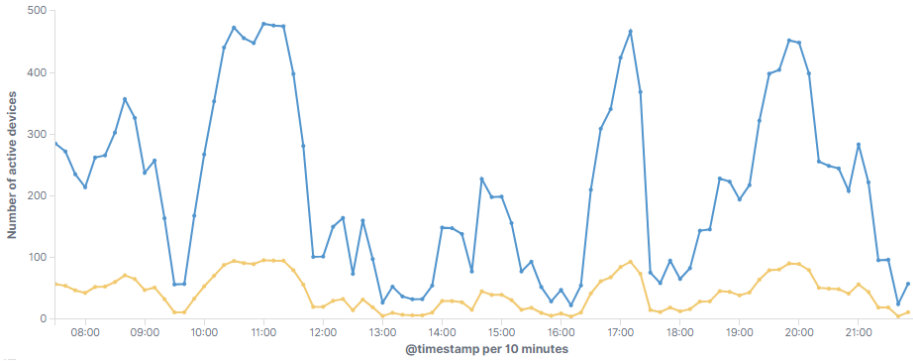

Fig. 4.22 Active devices (absolute number in blue, percentage in orange)

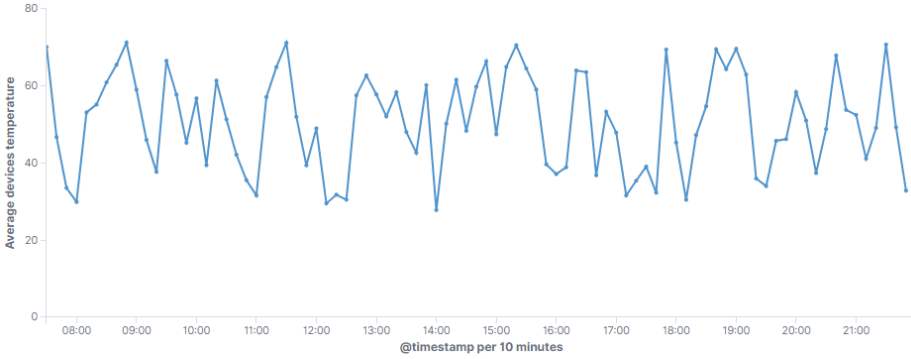

Fig. 4.23 Average temperature of the devices (ºC)

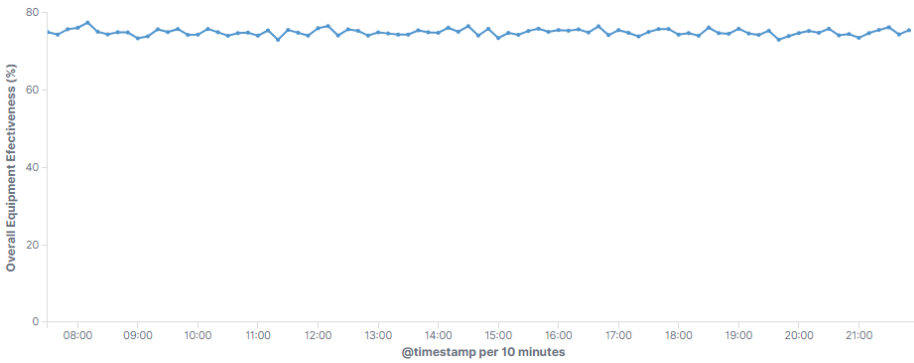

Fig. 4.24 Overall Equipment Efectiveness (OEE)

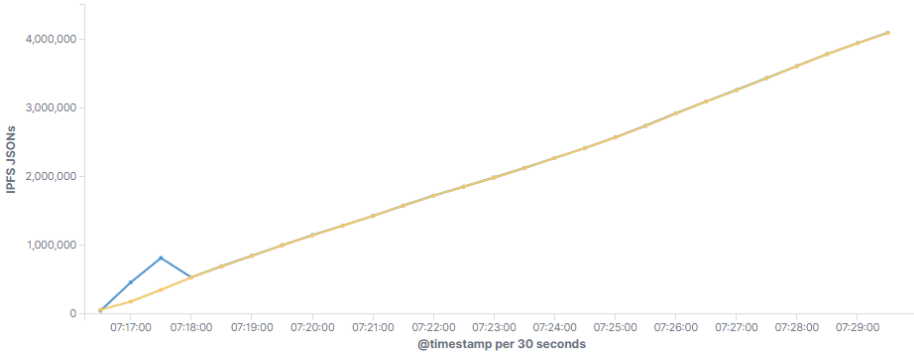

Fig. 4.25 Number of processed JSONs: IPFS (blue) and IOTA (orange)

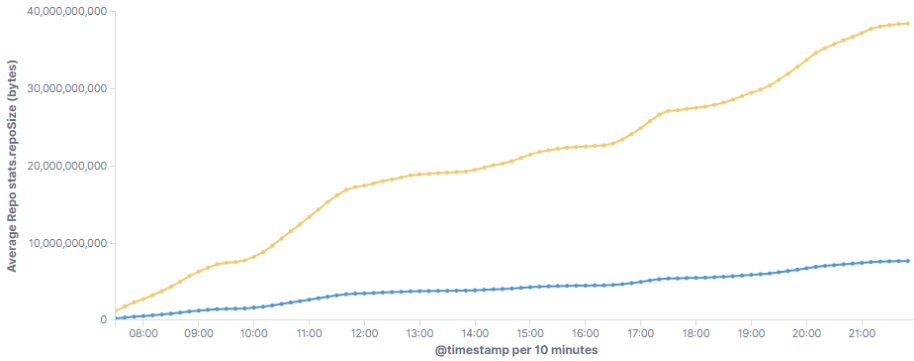

Fig. 4.26 Comparison in bytes between storage in IPFS (blue) and IOTA (orange)

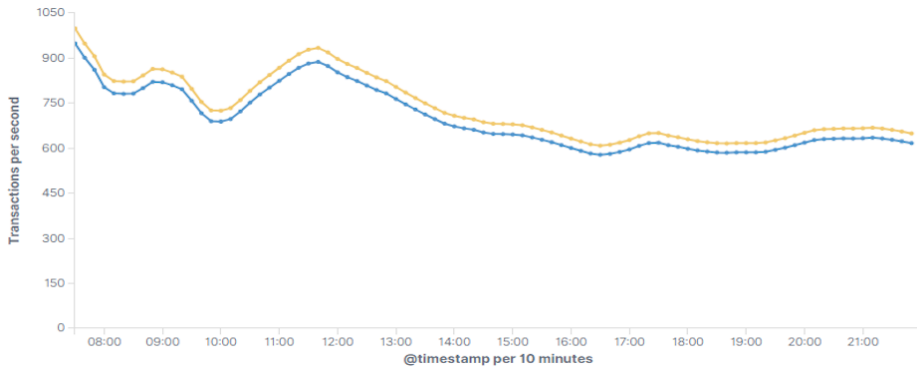

Fig. 4.27 Average DLTs throughput: IOTA (orange) and Polkadot (blue)

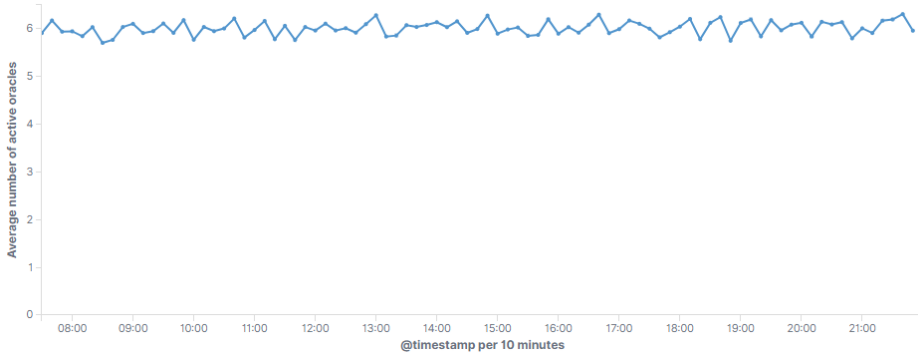

Fig. 4.28 Average number of active oracles

Results Discussion

Performance Analysis

In this work, decentralized oracles for data interoperability purposes were leveraged, i.e., to securely gather the external IIoT data model and perform a homogenization process of machine raw data. Decentralized oracle platforms such as ChainLink intend to enable the development of fast, decentralized, and secure oracles for different applications. ChainLink, however, is strongly linked to the Ethereum ecosystem. On the other hand, Polkadot, despite not being focused on the oracle services field of application, is a highly versatile and interoperable platform in which an oracle solution can be implemented apart from other conventional uses. Polkadot aims to achieve a high degree of interoperability to design a holistic DLT architecture for tomorrow's Industry 4.0.

However, despite the significant amount of security (i.e., data integrity) that a decentralized oracle mechanism brings to an architecture when providing data, some delays may be introduced due to the complexity of an additional decentralized network in between. Nonetheless, that would have been the case with ChainLink. By using Polkadot, the oracle platform can be integrated with the plant blockchain since Polkadot "parachains" have direct connection and compatibility through the main "relay" chain. Furthermore, the performance of Polakdot is significantly higher than other blockchains, such as Ethereum, on which ChainLink is currently based. Moreover, the direct connection between the oracle parachain and the interoperable plant blockchain relay chain incurs near-zero latency. Therefore, using a decentralized oracle service based on Polkadot for retrieving a JSON data model scheme does not have a significant impact on the performance of the scheme since, for each data model, only one request should be made. Finally, according to the measurements presented in Figure 4.28 from Section 4.7.2, on average, six oracles have been active for the given external data retrieval tasks. This number of oracles is appropriate to guarantee the complete decentralization of the architecture and almost instantaneously return the JSONs that comprise the Eclipse Unide data model.

As shown in the simulation results presented in Figure 4.27 from Section 4.7.2, in industrial environments, large amounts of data are generated, thus requiring significant processing and storage capacity. The presented monitoring system shows that the architecture is robust enough when handling great amounts of data. IOTA and Polkadot offer a great processing capacity, almost 1000 tps on average, which is sufficient in this type of environment. Even though Polkadot is not as fast as IOTA, this aspect is not relevant since the processing speed is most important where the data is generated. Furthermore, as shown in Figure 4.26 from Section 4.7.2, the use of IPFS greatly reduces the storage burden of the DLTs. In addition, the active devices measurements shown in Figure 4.22 from Section 4.7.2 prove that increasing the number of active devices does not incur a significant impact on performance.

The graphs generated from the continuous monitoring of the architecture help us to identify possible weak points in the process and, consequently, possible ways to improve the homogenization process, the data processing, as well as the management of possible costs. For example, the use of an oracle service could entail certain costs that should be optimized as much as possible by the companies. Thus, using the monitoring system, much more aspects could be analyzed and predicted , such as the incurred costs, the performance of the system, resource usage, device failures, etc. For example, in Figure 4.23 from Section 4.7.2, the average temperature of the devices is analyzed, where it can be seen that it has significant fluctuations within the range of 30 and 70 degrees ºC,

based on the intensity of the production process. Moreover, in Figure 4.24 from Section 4.7.2 the effectiveness of the industrial equipment (OEE) can be visualized, which gives us clues about the effectiveness of the machines. This information shows that the effectiveness of the machines is highly optimal during the entire simulated period, but with a certain margin of improvement.

Security Analysis

Regarding the security of the information, in the presented architecture, the integrity of the data is ensured during the whole process, from when the data is generated in production lines up until it is homogenized and finally exploited at the plant level. This is due to the use of secure DLT technologies throughout the whole process (i.e., production lines DAG DLTs, decentralized blockchain oracles for data homogenization, and plant processing blockchain). As shown in Figure 4.25 from Section 4.7.2, in the beginning, an attack in which great amounts of malicious data are generated is simulated. Nonetheless, the malicious data is finally discarded by the IOTA DLT. Such examples show that the monitoring of the architecture is also useful for visualizing possible cybersecurity attacks and other types of non-intentional incidents.

However, overall, the proposed architecture involves several components that may introduce potential security risks, including:

- IPFS. IPFS is a decentralized storage system, which means that it relies on a distributed network of nodes to store and retrieve data. While this can increase the availability and durability of the data, it also means that there is a risk that some nodes may not be trustworthy or may be compromised. To mitigate this risk, several security measures such as encryption and access control to ensure that only authorized parties can access the data stored in IPFS have been implemented.

- Decentralized oracles service. The proposed architecture involves using a decentralized oracle service to retrieve data models for the data homogenization process. This introduces a potential security risk, as oracle services are often centralized and may be subject to attacks or manipulation. To mitigate this risk, multiple oracle sources should be used, as well as implementing security measures such as cryptographic signing and verification to ensure the integrity and authenticity of the data retrieved from the oracle service. Another security issue of oracles might be the supply of unreliable information [271]. However, monitoring the oracles could help mitigate this issue. Thus, in this work, a monitoring system is already present.

- Interoperable plant blockchain. The interoperable plant blockchain is responsible for storing and managing smart plant homogenized data references and providing

access control to IPFS. To ensure the security of this blockchain, it is important to implement measures such as secure consensus algorithms, proper access control and permissions, and regular security audits. Additionally, measures such as encryption and secure communication protocols to protect the data stored on the blockchain are implemented.

- Smart contract-based notary scheme: The data exchange scheme involves using smart contracts to securely transfer data between the production lines DAGs and the plant blockchain. It is important to ensure that these smart contracts are properly tested and audited to ensure their security and correctness. Additionally, measures such as access control and permissions to ensure that only authorized parties can interact with the smart contracts are implemented.
- ELK-based monitoring. It is important to ensure that the ELK stack is properly configured and secured to protect against potential security risks and ensure the integrity and confidentiality of the data it processes. The latest version of the stack was used so that we can ensure that all the current known vulnerabilities have been mitigated.

Overall, it is important to ensure that all components of the proposed architecture are properly secured, and that appropriate measures are taken to mitigate potential security risks. This process involves implementing a combination of technical and organizational measures such as encryption, access control, cryptographic signing, security audits, and secure communication protocols.

Comparison with Other Solutions

The most similar DLT-based proposal is the architecture proposed by Jiang et. al [272]. This work presents a cross-chain framework for efficient and secure IoT data management using a consortium blockchain as the control station and other blockchain platforms customized for specific IoT scenarios as the backbone for IoT devices. The framework merges transactions based on a notary mechanism and is implemented using Hyperledger Fabric and IOTA. However, this work shows a much lower throughput capacity (600 tps vs. 900 tps), and higher overall latency. Furthermore, the security robustness of the aforementioned architecture is not clear, since the authors tackle security concerns only by designing a simple access control system. Moreover, this work goes one step further and perform industrial data homogenization and exploitation instead of focusing exclusively on simple data transfer between DLTs. Finally, advanced monitoring of the whole scheme by using the ELK stack is also provided.

However, an industrial data processing, monitoring, and homogenization process can also be non-DLT based. In fact, nowadays, an overwhelming number of real-world

industrial architectures are non-DLT based, since this technology is relatively new, and industrial processes take a considerable time to incorporate new technologies. However, here are some potential alternatives to DLTs that could be used for efficient and secure data management and homogenization in Industry 4.0:

- Centralized databases: A centralized database is a single repository of data that is managed and maintained by a single entity. This can be an efficient way to manage data in the IoT, as it allows for quick and easy access to data and can scale to handle large volumes of data. However, it can also be vulnerable to security threats, as a single point of failure can compromise the entire system. Furthermore, centralized databases could have serious bottlenecks and collapse in the face of a large amount of data that needs to be processed and homogenized.
- P2P networks: P2P networks allow devices to communicate directly with each other without the need for a central server or authority. This can be an effective way to manage data in IoT, as it allows for decentralized control and can be highly scalable. However, it can also be less secure, as it relies on the security and reliability of individual devices, and the lack of a robust consensus and data blocks cryptography links, as is the case of the most used DLTs.
- Cloud-based solutions: Cloud-based solutions allow data to be stored and accessed on remote servers, which can be accessed over the internet. This can be a convenient and scalable way to manage data in IIoT, as it allows for easy access to data from any location. However, it can also be less secure, as data is stored on servers that may not be physically secure. Furthermore, cloud storage usually entails much higher economic costs than DLTs, especially compared to the more advanced solutions such as IOTA, which does not require fees, or Polkadot, whose fees are low or even zero in private networks.

Thus, this architecture not only ensures data integrity and security at every stage of the process but also delivers high performance for handling large amounts of IIoT data. Additionally, it is designed to be cost-effective, making it an attractive solution for businesses looking to leverage the benefits of IIoT with relatively low monetary costs. Furthermore, the implemented monitoring system also provides comprehensive real-time analysis, threat detection, and optimization suggestions across the whole process.

4.7.3 Business Layer and Whole Architecture

In this subsection the implementation process of the "Business DLT Layer" and the entire architecture as a whole is described. The third layer of the architecture is built on top

of Hyperledger Fabric (v2.2), which is an open-source, consortium oriented blockchain that has smart contract capabilities, private channels, customizable consensus and zero fees. This platform is widely used by companies, researchers and developers to create efficient consortium blockchains for a wide range of applications [273].

Implementing a Hyperledger Fabric blockchain with private channels, oracles for trustworthy external data retrieval, smart contracts (chaincodes), and a custom developed connector for interoperability purposes involves several steps. First, a network that includes the required number of organizations, peers, and channels is designed. Then, the oracle service is established to allow the overall blockchain architecture and its chaincodes to interact with external data sources when needed. Furthermore, private channels need be created in order to limit the access of private data that is sent between the organizations. The custom connector that was defined in the previous section to enable interoperability with other blockchains is also implemented. Finally, the blockchain network is deployed on a set of physical or virtual machines and tested in a development environment before being rolled out to production. Once the network is live, its privacy, security and performance capabilities are monitored in order to validate the effectivity of the solution and detect possible failures and improvement opportunities.

The implementation has been carried out on a laptop with Ubuntu operating system, an i7 processor, 16 GB of RAM and a 512 GB SSD. As for end users to interact with the network, the requirements are significantly lower as they would interact with the blockchain through applications and interfaces that abstract away most of the technical complexities of the network itself.

Use case 1 - Fagor Automation machines monitoring

A real-world use case resulting from a partnership between the IKERLAN research center[6] and Fagor Automation[7], a worldwide leading industrial company with extensive experience in the development and manufacture of products for the automation and control of machines is implemented.

This use case aims to monitor three parameters from the Computer Numerical Control (CNC) machines that belong to Fagor Automation: installed software, validation code and parametrization settings. All illegal changes are registered and processed throughout the industrial plants, and then shared to the proposed business blockchain within a smart contract, that establishes the status of the warranty based on a cumulative score given by the level of severity of the detected changes. In order to determine

[6] https://www.ikerlan.es/
[7] https://www.fagorautomation.com/

whether an installed software is critical or not, external "critical software" list from the machines vendor using trustworthy blockchain oracles is retrieved. For this task the official Fagor Automation CNC simulator[8] is set up in the evaluation machine.

Figure 4.29 shows the implementation of the script that monitors the industrial machines and detect suspicious changes related to the three aspects that are mentioned above: installed software, validation code and parametrization settings. When changes are detected, the machine data and the generated changes log are processed through the production lines DLT, and then sent to the industrial plant, where severity is calculated.

The process begins with an interaction with the blockchain oracles, which provide meaningful data to the smart contracts (chaincodes). The oracles provide a "critical software list", which contains various software that are deemed critical for the machine's functionality and integrity. Since the critical software list is external, it can be modified only by the CNC owner - in this case, Fagor Automation.

The score limit is set at 10. The scoring is calculated based on several factors:

- Trivial Changes (Score: 1): These changes include modifications in system files or processes that are not deemed critical for the operation of the machine, non-critical software updates, or slight tweaks in parametrization settings. These are changes that won't significantly affect the machine's performance or compromise its security.

- Moderate Changes (Score: 5): This category includes changes that can potentially affect the machine's performance but are not likely to compromise its security. For instance, changes in the validation code that do not affect critical processes, installation of software that is not part of the critical software list, or moderate adjustments in parametrization settings.

- Severe Changes (Score: 10): These are critical changes that could potentially lead to a breach of the machine's warranty terms. Examples include the installation or update of critical software without proper validation, alterations in the validation code that affect critical processes, or substantial modifications in parametrization settings that can drastically impact the machine's performance.

The score considers the nature of the change: if there is a modification, deletion, or addition of critical system files or processes, it could attract a higher score. The number of changes is also a factor: if numerous alterations occur in a short time span, the score increases. Furthermore, the score is influenced by the nature of the software itself: changes in some software may be more detrimental than others, based on its role

[8] `https://hmielite.fagorautomation.com/`

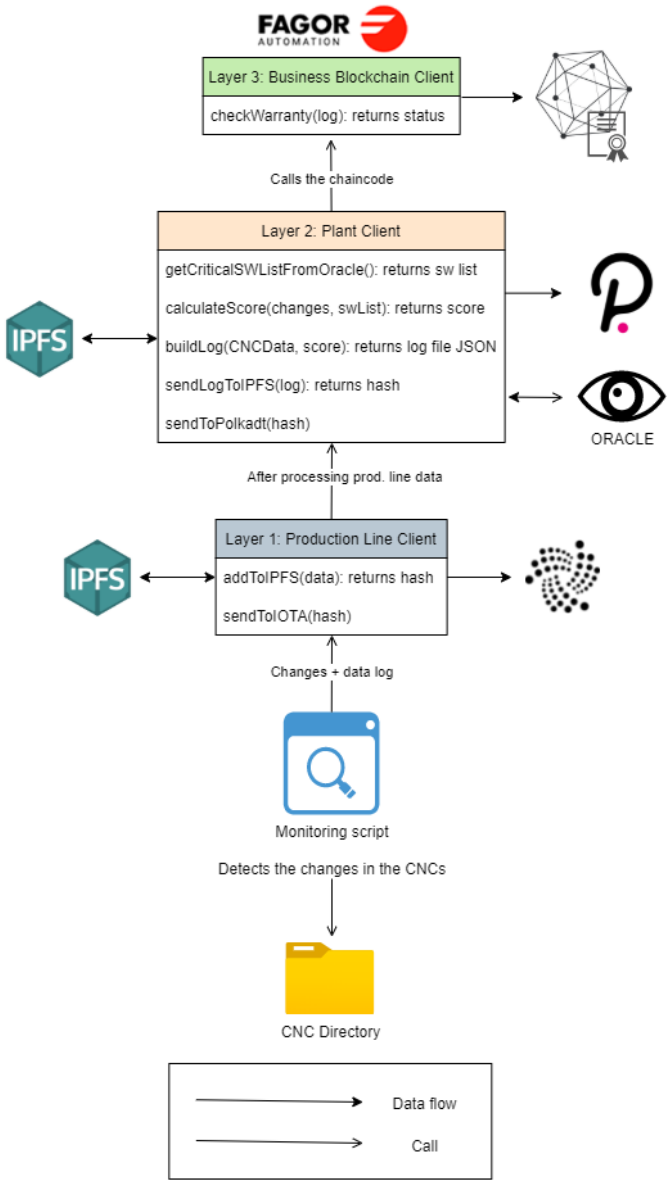

Fig. 4.29 Fagor Automation use case diagram

and importance within the system. It is important to notice that in specific cases, such as repeated install and uninstall of the same software, the score would be altered once.

After detecting and scoring these changes, a log is generated for each machine at the plant level. This log contains information about the changes and the identity of the machine, and is shared with the Fabric business consortium blockchain. An example

log is shown in Listing 9. After sharing the information from this log with the business consortium blockchain, the chaincodes allow the manufacturer of the machines (Fagor Automation) to check the status of the warranty. If the changes detected have been assessed as illegal and critical by exceeding the established score threshold, the chaincode triggers the revocation of the warranty.

Listing 9 Generated log JSON from industrial plant example

```
{
  "alterationDetails": "Unknown software alteration",
  "score": 1,
  "totalScore": 2,
  "CNCData": {
    "machineId": 8352,
    "machineModel": "Model-15",
    "machineType": "Milling",
    "machineStatus": "Under Maintenance",
    "plant": "Plant-C",
    "productionLine": "Line-19",
    "location": "Arrasate"
  }
}
```

In addition to these features, the chaincode also contains the ability to handle exception scenarios. For instance, if an unanticipated modification takes place that has not been defined within the existing rules, the chaincode does not ignore or reject it outright. Instead, it flags it as an "unknown alteration" and triggers an alert for manual investigation. This ensures that the system remains flexible and adaptable to unexpected changes, thus, maximizing the overall reliability of the warranty management system.

Furthermore, the chaincode is designed to be continually updated and optimized based on the evolving industrial environment and warranty policy revisions. Its modular architecture allows new rules to be added or existing ones to be adjusted without affecting the overall system performance or requiring a complete overhaul of the chaincode.

Lastly, the chaincode provides an audit trail that ensures compliance with regulatory requirements. Every action taken by the chaincode is timestamped and logged, providing a detailed, immutable record of all decisions related to warranty scoring and revocation. This not only supports compliance but also aids in dispute resolution, as it provides irrefutable evidence of the events leading to the warranty's revocation.

The extended functionalities of the chaincode contribute to a robust and efficient warranty management system. Algorithm 1 demonstrates how the chaincode checks the warranty status using generated logs from industrial plants. Each step of the algorithm encapsulates the sophisticated mechanisms incorporated in the chaincode design, ensuring precision and reliability in managing complex warranty scenarios.

Algorithm 1 Check Warranty Based on Score Chaincode

```
1:  procedure CHECKWARRANTY(jsonPayload)
2:      warrantyDataJSON ← jsonPayload
3:      if warrantyDataJSON = null then
4:          return Error : εInvalidJSONpayloadε
5:      else
6:          warrantyData ← UNMARSHAL(warrantyDataJSON)
7:      end if
8:      totalScore ← warrantyData.TotalScore
9:      if totalScore >= 10 then
10:         warrantyStatus ← "Warranty is void"
11:     else
12:         warrantyStatus ← "Warranty is valid"
13:     end if
14:     return warrantyStatus
15: end procedure
```

Use case 2 - Machine price adjustment based on OEE

An use case based on the OEE profitability calculations[9] is implemented, which consists of an industrial plant sharing information about its machines' effectiveness rates with the manufacturers so that they can adjust the renting prices. This use case is leveraged as a default dummy use case, aimed primarily at providing an initial validation of the architecture being developed. This trial serves to highlight the key functionalities and capabilities of the system, ensuring they are performing as expected while demonstrating their value in a practical context.

This particular use case employs the principles of OEE, an essential component in modern manufacturing operations. The OEE model encapsulates the understanding and evaluation of how effectively a manufacturing operation is utilized. The factors measured by OEE are fundamental elements in identifying losses, benchmarking progress, and improving the productivity of manufacturing equipment (i.e., availability, performance, and quality).

Within this sphere, using OEE is crucial for industries to ensure optimum productivity, performance quality, and machine availability. It offers valuable insights for the stakeholders, such as industrial plant operators and equipment manufacturers, facilitating informed decision-making processes like adjusting renting prices based on the machinery's effectiveness rates.

In essence, this use case is more than a mere functional verification step. It also offers a practical, industry-relevant example demonstrating the potential and versatility

[9] https://www.leanmap.com/calculator/equipment-productivity/

of the developed architecture. Through its implementation, other potential and more realistic applications and use cases will be spurred, further exploiting the capabilities of the system.

As shown in previous works [274] [275], the machine information from the industrial plant is stored in an IPFS decentralized file in collaboration with a Polkadot interoperable blockchain, that stores the references of the data. Thus, the IPFS + Polkadot blockchain is connected to a custom Hyperledger Fabric chaincode using the Polakdot{.js}, the IPFS NodeJS API and the Go and JavaScript Fabric Software Development Kit (SDK).

Regarding the oracles, in this case, a blockchain oracle service based also on Polkadot is set up, since Polkadot offers immense interoperability capacities. Thus, the process of interacting with the oracles for retrieving external data will be similar to the interaction process between the Polkadot + IPFS plant blockchain and the Hyperledger Fabric blockchain that was described before.

In the middle of the diagram from Figure 4.30, we have the Hyperledger Fabric blockchain, which communicates several business partners through smart contracts via private channels. Furthermore, blockchain oracles are set to develop secure external data to the agreements that are being executed in the network. Finally, industrial enterprises connect their plant blockchains using interoperability adaptors in order to share information with their partners.

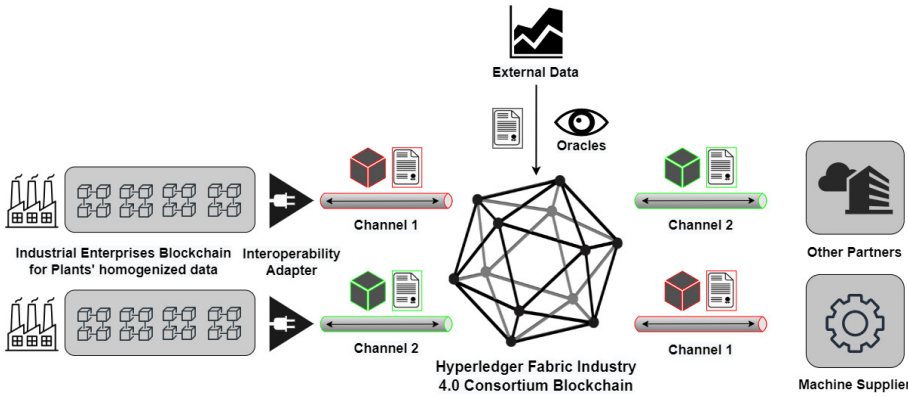

Fig. 4.30 Hyperledger Fabric blockchain implementation diagram

The parameters that should be extracted from the industrial data so that the industrial plant can provide meaningful information to the machine providers using a Go chaincode are defined. Specifically, several values are gathered, such as the appraised value of the machine, the manufacturer's name, the ID of the machine, its current owner, and its OEE, which indicates the production effectiveness percentage of each machine.

Figure 4.31 shows the actual structure of a Hyperledger Fabric chaincode with a register of seven machines. In Listing 10, the example JSON structure of "machine1" is shown in detail.

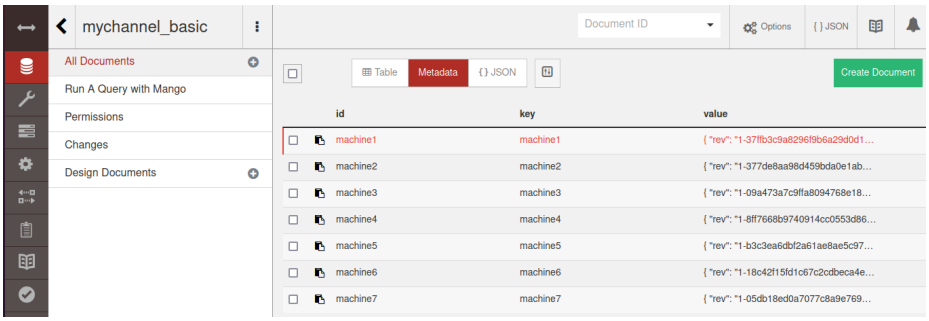

Fig. 4.31 Hyperledger Fabric chaincode machines list

Listing 10 "machine1" JSON structure example

```
{
  "_id": "machine1",
  "_rev": "1-37ffb3c9a8296f9b6a29d0d1f4b774db",
  "AppraisedValue": 30000,
  "ID": "machine1",
  "Manufacturer": "m1",
  "Oee": 95,
  "Owner": "CompanyX",
  "~version": "CgMBBgA="
}
```

The machine information is retrieved from the IPFS storage using the hash that is stored in the plant Polkadot blockchain. This task is achieved by employing the Polakdot{.js} API and the IPFS NodeJS API. Specifically, firstly the architecture connects to the Polkadot blockchain, from which the hash key of the data is retrieved using the machine ID and the timestamp. Then, using the aforementioned hash, the actual data is retrieved from the IPFS API. Finally, the data is processed so meaningful information can be obtained for the chaincodes.

The pseudocode in Algorithm 2 shows the procedure that is used to feed the Hyperledger Fabric chaincode with the machine information from the Polkadot plant blockchain, thus achieving the desired DLT interoperability. From lines 3 to 7 the gateway the performs the connection to the Fabric blockchain and its private channel is defined. Then, the SDK is employed to interact with the chaincode by, for example, adding a new register of data. We can also perform more actions such as update, delete or simply read it.

The Go chaincode that is deployed to evaluate the machine's OEE allows us to

Algorithm 2 Send the plant data to Hyperledger

```
1:  procedure SENDDATATOHYPERLEDGER(machineData)
2:      // Connect to the Hyperledger Fabric network
3:      gateway ← new Gateway()
4:      gateway.connect(CCP, wallet, org1UserId, discovery)
5:      network ← gateway.getNetwork('channel')
6:      contract ← network.getContract('chaincode')
7:      contract.submitTransaction()
8:      // Submit the machine information for 'machine1'
9:      contract.submitTransaction('CreateAsset', machineData.id, machine-
    Data.manufacturer, machineData.oee, machineData.owner, machine-
    Data.appraisedValue)
10:     // Update the machine information for 'machine1'
11:     contract.submitTransaction('UpdateAsset', 'machine1', 'm1', '95', 'Compa-
    nyX', '28000')
12:     gateway.disconnect()
13: end procedure
```

adjust the renting price of the machine based on its performance.

The performance of the machine is categorized into five tiers: "Poor" "Below average" "Average" "Good" and "Excellent". These categories are derived from the machine's OEE values, which are calculated using the formula given in Belohlavek's work on OEE [276].

The categorization ranges are as follows:

- "Poor" performance: Any machine with an OEE less than 60 falls under this category. Such low OEE scores often point to significant equipment failures and poor quality control.
- "Below average" performance: If a machine's OEE lies between 60 and 74 (inclusive), it is deemed to perform below average. Machines in this category may experience intermittent issues affecting their efficiency.
- "Average" performance: A machine with an OEE between 75 and 89 (inclusive) is said to have average performance. Machines in this range are generally functioning as expected but may benefit from tweaks to enhance efficiency.
- "Good" performance: When a machine's OEE lies between 90 and 94 (inclusive). Machines in this category have good efficiency and minimal downtime.
- "Excellent" performance: Any machine with an OEE of 95 or above is categorized as excellent. They produce high-quality output consistently and serve as industry benchmarks for performance.

Once categorized, the respective OEE values are registered in the chaincode. As the data is encoded in the blockchain, it is tamper-proof and can only be altered by modifying and redeploying the entire chaincode.

Rent prices are adjusted according to the OEE tier. A machine with a "Poor" performance rating sees a decrease in rent price by 30%, while a "Below average" performance rating implies a 10% reduction. A performance rating of "Good" prompts a 1% increase in the rental price, and an "Excellent" rating prompts a 2% increase.

The adjusted rent prices reflect the efficiency and performance of each machine, providing a transparent and equitable pricing model for all parties involved.

To illustrate, the pseudo code for the chaincode function, "AdjustRentPriceBasedOnPerformance", is provided in Algorithm 3.

Algorithm 3 Adjust Rent Price Based on Performance

```
1:  procedure AdjustRentPriceBasedOnPerformance(id)
2:      asset ← ReadAsset(id)
3:      rentPriceJSON ← GetState(RentPrice_id)
4:      if rentPriceJSON = null then
5:          rentPrice ← 1000                          ▷ Default rent price
6:      else
7:          rentPrice ← Unmarshal(rentPriceJSON)
8:      end if
9:      oee ← asset.Oee
10:     if oee < 60 then
11:         performance ← "Poor"
12:         rentPrice ← rentPrice − rentPrice ∗ 0.30
13:     else if oee < 75 then
14:         performance ← "Below average"
15:         rentPrice ← rentPrice − rentPrice ∗ 0.10
16:     else if oee < 90 then
17:         performance ← "Average"
18:     else if oee < 95 then
19:         performance ← "Good"
20:         rentPrice ← rentPrice + rentPrice ∗ 0.01
21:     else
22:         performance ← "Excellent"
23:         rentPrice ← rentPrice + rentPrice ∗ 0.02
24:     end if
25:     rentPriceJSON ← Encode(rentPrice)
26:     PutState("RentPrice_ε + id, rentPriceJSON)
27:     return performance
28: end procedure
```

It retrieves the machine's details, calculates the OEE, adjusts the rent price based

on the OEE category, and updates the machine's state with the new rent price in the blockchain. This Go chaincode implementation encourages machine owners to improve and maintain their machine performance, while at the same time, giving renters a clear understanding of what they are paying for. The robustness of this implementation relies on the immutable nature of the blockchain, ensuring the integrity and reliability of the recorded OEE values and rent prices.

Discussion

In the ensuing discussion section, the potential of the proposed solution to revolutionize business processes within Industry 4.0 is explored. Also, the analysis and implementation of the proposed platform in a realistic industrial case scenario is discussed. Furthermore, the importance of data integrity and traceability in achieving a holistic DLT architecture and analyze the security and performance metrics of the implementation is emphasized.

Security analysis

This subsection discussed the threats to validity in terms of security within the implementation of an Industry 4.0 smart contract-based consortium blockchain with Hyperledger Fabric that interacts with external plant blockchains (Polkadot) located in each industrial plant via gateway API connections (Polkadot{.js}) and Fabric chaincodes. The complex nature of this system presents multiple challenges that might arise. By examining each component, their interactions, and potential vulnerabilities, the aim is to provide a comprehensive understanding of the threats that must be addressed to maintain the system's overall reliability.

To identify the possible threats, the process begins by firstly dissecting the components and their interactions, such as the Hyperledger Fabric blockchain, the smart contracts (chaincodes in Fabric), the Polkadot blockchains and the gateway connections for interoperability. Then the communication and data sharing between components is assessed to identify possible attack vectors and scrutinize external dependencies for potential vulnerabilities.

Below is provided a detailed list of possible security threats of the architecture and how these risks are mitigated:

- Tampering: The system utilizes cryptographic hashing functions and digital signatures to ensure data integrity. Both Polkadot and Hyperledger Fabric blockchains implement secure consensus algorithms and permissioned network characteristics. Additionally, data is stored in an append-only, distributed ledger, making unauthorized data tampering nearly impossible.

- Repudiation: The use of digital signatures and an immutable, append-only distributed ledger ensures that all transactions and actions are attributable to their respective users or nodes. This design eliminates the possibility of repudiation, providing a reliable audit trail.
- Information Disclosure: Data confidentiality is preserved through encryption, both at rest and in transit. All communication between nodes, as well as between the Polkadot and Hyperledger Fabric blockchains, is secured using TLS. Access control mechanisms are also in place to ensure that only authorized users and nodes can access sensitive data.
- Denial of Service: The distributed nature of the system provides inherent resilience against Denial of Service (DoS) attacks. Both Polkadot and Hyperledger Fabric blockchains implement measures to prevent or mitigate the impact of such attacks, including rate limiting, transaction validation, and blacklisting of malicious nodes.
- Smart contract vulnerabilities: Smart contracts are developed using secure coding practices and undergo rigorous testing, including unit testing, integration testing, and formal verification, to minimize the risk of vulnerabilities.
- Insecure APIs: The gateway API connections between the Polkadot and Hyperledger Fabric blockchains follow industry best practices in API security, such as proper authentication, input validation, and rate limiting.
- Consensus mechanism attacks: Both Polkadot and Hyperledger Fabric have secure consensus algorithms that are designed to resist attacks such as Sybil and Eclipse. The use of a consortium blockchain, where membership is controlled, further reduces the risk of consensus-related attacks, as malicious nodes are less likely to gain a majority control of the network.
- External dependency vulnerabilities: The system's dependencies, such as the Polkadot{.js}, are carefully vetted and monitored for security vulnerabilities. Regular updates and patches are applied to minimize the risk posed by these external components.

Performance analysis

In this subsection, the performance analysis of the Hyperledger Fabric implementation using Hyperledger Caliper[10], a blockchain benchmarking tool, is presented. Caliper

[10] `https://www.hyperledger.org/use/caliper`

was selected due to its ability to produce a set of comprehensive performance reports and its direct compatibility with Hyperledger Fabric.

The performance evaluation framework aims to measure various important metrics including transaction throughput, latency, resource usage, and chaincode execution time. Hyperledger Caliper is configured to simulate a variety of workloads representing different transaction sizes and volumes. This is done to evaluate the performance of the implementation under different operating conditions.

The tests were run multiple times under each scenario to ensure the robustness and consistency of the results. The data recorded by Caliper is then analyzed for each metric under the different scenarios. Transaction throughput and latency are analyzed to understand the capacity of the network and the responsiveness of transactions, respectively. Resource usage offered insights into the computational efficiency of the implementation, while chaincode execution time analyzes the efficiency of the custom chaincodes.

The evaluation of the Hyperledger Fabric implementation yielded significant insights into the performance of the network. In terms of transaction throughput, the system was able to process an average of 106 transactions per second in each scenario. A slight decrease to 94 transactions per second under high workload conditions was observed, indicating robust scalability. Figure 4.32 shows the evolution of the throughput in 1000 transactions.

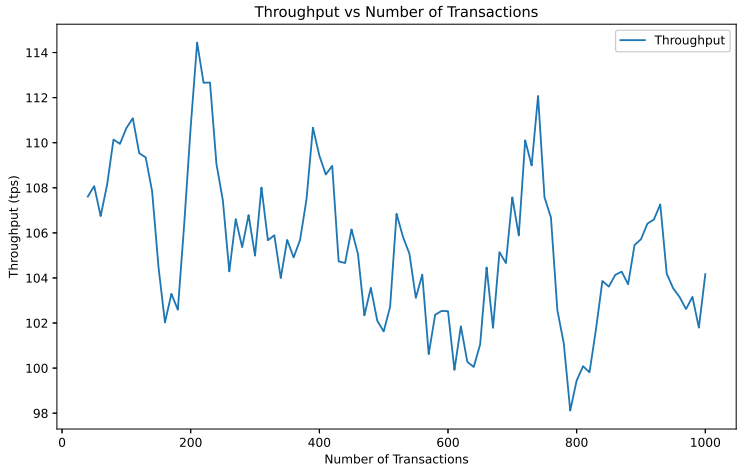

Fig. 4.32 Throughput (TPS) evolution in 1000 transactions

Latency, which is crucial for the responsiveness of transactions, averaged at 1.23 seconds. In high workload scenarios, it increased marginally to 2 seconds, but remained

within acceptable limits.

Resource usage was optimal throughout the tests. The CPU usage peaked at 25% under high workloads, while memory utilization never exceeded 84 MB of usage. This underlines the computational efficiency of the implementation.

Lastly, the chaincode execution time was analyzed. The execution time for the Check Warranty Based on Score chaincode was on average 0.14 milliseconds, proving its efficiency in deciding the warranty status based on the total score.

Since the typical performance metrics for business consortium networks consist of several dozens to hundreds transactions per seconds and latencies of approximately 1 to 5 seconds [277], the results confirm the effectiveness and efficiency of the Hyperledger Fabric implementation, demonstrating its readiness for larger-scale application.

4.8 Summary and Conclusion

In this chapter, a comprehensive multi-layered DLT architecture tailored specifically for Industry 4.0 was presented. The architecture is composed of the following parts:

Layer 1 - Data Source Layer: At the core of the proposed architecture is the data source layer. This layer, designed with a DAG DLT, is optimized for the machine and production line levels in the modern industrial landscape. Recognizing the limitations of lightweight devices in IIoT environments, the design introduces several critical improvements. These enhancements encompass the inclusion of lightweight devices, advancements in cryptography, storage optimization, and refinement of consensus mechanisms. As a result, a notable reduction in energy consumption and computational burden is witnessed, paving the way for a scalable, efficient, and adaptable solution that meets the diversified demands of contemporary industrial settings.

Layer 2 - Bridge Layer: Transitioning to the next layer, a homogenization process that ensures data consistency and integrity across the IIoT network was designed. By employing decentralized oracles and integrating with technologies such as the Polkadot interoperable blockchain and the IPFS decentralized storage, the data integrity is guaranteed throughout its lifecycle within an industrial setting. Complementing this is a monitoring system that visually represents the entire process, further enhancing transparency and assurance for stakeholders.

Business Blockchain Decision Tree: Preceding the third layer, a robust decision tree to streamline the selection of a suitable DLT platform for business-focused applications was introduced. This decision tree, a product of an exhaustive analysis of the sector's needs, is intended to minimize conflicts and offer clarity. Through this, stakeholders can easily navigate the complex landscape of DLT platforms and make informed decisions.

Layer 3 - Business Layer: Lastly, in the third layer, the work delves deep into the potential and challenges of integrating blockchain and smart contract technologies into Industry 4.0. The proposed architecture revolutionizes business process management by countering issues such as centralization, traceability gaps, and automation deficiencies. By harnessing the capabilities of Hyperledger Fabric, a system that emphasizes privacy, security, and automation was developed. The effectiveness of this solution is underscored by the success in executing automatic agreements via smart contract technology, with ensured compatibility, performance, and affordability.

In conclusion, this chapter provides a holistic, multi-layered approach to integrating DLT in Industry 4.0, addressing its unique challenges and leveraging its opportunities. Through rigorous design and real-world validation, the way for a more resilient, efficient, and transparent industrial future was paved.

Chapter 5

Concluding remarks

5.1 Thesis Conclusions

The beginning of this dissertation defined the main objective of the thesis as "*To study DLTs in depth, with a view to design a holistic DLT architecture that covers the whole cycle of the data (from when it is generated from the IoT machines up until it is exploited for business purposes) and addresses the aforementioned Industry 4.0 challenges without neglecting the particular challenges and requirements of the actual DLT technologies*".

The goal is linked to the lack of a de-facto solution to achieve this task, and the dissertation has provided a way to bridge this gap. Overall, this thesis presents the design of a multi-layered DLT based architecture that covers the whole data lifecycle, from when they are generated up until they are processed and exploited for business purposes. More specifically, these solutions are contextualized in the resolution of two research questions, the answers and analysis of which are provided below:

Question 1: *Is the automation pyramid the ideal model on which we can design an architecture based on DLTs for Industry 4.0?* - Answered in Chapter 2

This thesis undertook an in-depth exploration of Industry 4.0 transformative effect on the manufacturing industry, which also provided insights into methods that can be employed to accurately delineate the structure of an Industry 4.0 ecosystem. The shift towards Industry 4.0 is indicative of a paradigm shift towards amplifying the efficiency, productivity, and competitiveness of industrial enterprises by leveraging advanced digital technologies, exploiting data-driven insights, and creating interconnected systems. One of the most significant tools in understanding this transformative shift is the automation pyramid, a model that signifies the hierarchical structure of contemporary factories. It outlines a methodical approach for understanding and deploying the assorted strata of technology, data flow, and communication protocols that are embedded within

the manufacturing process. This crucial tool allows enterprises to successfully navigate their transition towards Industry 4.0.

This chapter notably introduced an adaptation of the automation pyramid into a four-layered industrial scenario that serves as the context for this research. This adaptive model facilitates a structured procedure for Industry 4.0 adoption, wherein each layer — the machine, the production line, the plant, and the consortium — caters to discrete aspects of the industrial process. From optimizing individual assets and whole production lines to the integration of several plants into a single business consortium, this model sets forth a comprehensive strategy for applying Industry 4.0 principles.

Furthermore, by integrating this structured approach with the application of DLTs, it can be expected to enhance the data lifecycle management across the entire Industry 4.0 ecosystem. By offering a transparent, secure, and decentralized approach to data management, DLTs can streamline data exchanges, promote data integrity, and ensure data privacy in the increasingly interconnected Industry 4.0 ecosystem.

As a consequence of the answer to Question 1, the need of examining the current state-of-the-art (Chapter 3) in DLT architectures for many relevant fields focusing especially on the Industry 4.0 was identified.

This thesis conducted a detailed exploration and analysis of 126 blockchain architecture proposals. The aforementioned proposals, cover an impressive range of sectors, providing a broad understanding of the unique features, components, and evaluations of the architectures.

The conducted review discerned a rapidly growing interest in DLT-based solutions, with research particularly focusing on challenges associated with deploying these technologies in resource-constrained environments. This suggests that DLT architectures can be instrumental in effectively dealing with such unique constraints.

While there is a profusion of proposed solutions in this sphere, it became increasingly evident through the study that there is a pressing need for additional research, particularly in maintaining a delicate balance between security and efficiency within the DLTs. The demand for enhancing blockchain-based IoT architectures, especially regarding interoperable solutions functioning across various platforms and blockchains, is unmistakable.

Furthermore, DLT interoperability, the deployment of smart contracts, and the traceability inherently offered by this technology are significant areas of potential in automating and streamlining processes across multiple sectors. The ability to trace transactions can be especially beneficial in ensuring data integrity and transparency, critical to various applications, including data management in Industry 4.0.

Furthermore, the findings of this study shed light on DAG DLTs as an intriguing yet

largely untapped terrain within the DLT landscape. These structures, with their lower energy consumption, absence of fees, and high throughput, pose an appealing alternative. Nevertheless, the relevance of traditional blockchains persists, particularly when coupled with round-robin or vote-based consensus algorithms and layered Edge architectures, which prove efficient across various IoT scenarios. Other promising solutions include easing the storage burden of the blockchain through decentralized databases such as IPFS and developing post-quantum cryptographic solutions to mitigate potential quantum computing threats to existing blockchain architectures.

In light of these findings, light is shed on the improvement and application of DLT technologies within IoT-oriented fields, including Industry 4.0. This perspective is rooted in the potential these technologies demonstrate in enhancing security, ensuring data immutability, and improving traceability. This thesis carried these insights forward as it strove to design a comprehensive DLT architecture for Industry 4.0, leveraging multiple DLT architectures to develop a secure, efficient, and transparent digital infrastructure.

Therefore, now the core research question can be answered, which would fulfill the objective of the thesis:

Question 2: *How can we design an architecture for Industry 4.0 based on DLTs that satisfies the above requirements?* - Answered in Chapter 4

This thesis proposes a multi-layer DLT architecture tailored to meet the specific requirements of data management and security within the Industry 4.0 scenario outlined in Chapter 2. The virtues of DLT technologies as secure, decentralized, and tamper-proof methods for storing and sharing data make them ideal for overcoming data management challenges associated with modern industrial processes.

The design of this multi-layer DLT architecture is rooted in the goal of enabling secure and efficient data handling across the complete industrial ecosystem. This covers every stage, from the machine level to the consortium level. The architecture's integration of DLT networks at various stages promotes seamless communication and information flow, fostering a connected, agile, and secure manufacturing environment. This comprehensive approach encourages trust and collaboration among stakeholders, aids the integration and interoperability of industrial processes, and opens avenues for implementing advanced Industry 4.0 solutions, such as smart contracts and decentralized applications.

The architecture's layers are as follows:

- Layer 1 - Data source DLT layer: Positioned at the production line level, this layer captures, stores, and manages data generated by IIoT devices in real-time, ensuring data integrity and providing a tamper-proof record of machine-level ac-

tivities essential for traceability and auditability. This layer facilitates real-time monitoring and control of individual machines and their interconnections, enabling efficient production line management and the identification of potential issues.

- Layer 2 - Bridge DLT layer: At the plant level, this layer aggregates data from the first layer production lines, homogenizes it for further processing, and securely transmits it to the business DLT layer. Thus, it promotes optimal resource allocation and coordination among various production processes. It supports the integration of multiple plants and production lines, fostering a connected and efficient manufacturing ecosystem. It also enables the implementation of plant-wide performance monitoring and analytics, enhancing operational efficiency and competitiveness.

- Layer 3 - Business DLT layer: Operating at the consortium level, this layer processes, analyzes, and manages data from multiple plants to support strategic business decisions and derive valuable insights. It ensures data security, privacy, and traceability, allowing consortium members and external stakeholders to establish trust in the system. This layer also facilitates the implementation of smart contracts for automating business processes, such as supply chain management, product tracking, quality control, and regulatory compliance.

The proposed multi-layer DLT architecture represents a significant leap towards achieving the main objective of the thesis: designing a holistic DLT architecture for Industry 4.0 that effectively addresses the primary challenges in cybersecurity, data standardization, system interoperability, scalability, automation and costs:

- Addressing Cybersecurity: Each layer of the architecture plays a vital role in maintaining a secure environment for data. From the data source DLT layer, which ensures real-time data integrity and provides a tamper-proof record of machine-level activities, to the business DLT layer, which guarantees data security, privacy, and traceability. Thus, the multi-layer architecture offers an inherently secure and tamper-resistant environment that preserves data privacy and integrity.

- Promoting Data Standardization: The second layer, the bridge DLT layer, is instrumental in aggregating, homogenizing, and transmitting data to higher layers. It plays a significant role in ensuring efficient data standardization, promoting seamless communication and data exchange between different production lines within a plant.

- Enhancing System Interoperability: The proposed architecture, by integrating DLT networks at various stages of the industrial process, enables seamless information flow and communication. This promotes system interoperability, fostering a connected and efficient manufacturing ecosystem, thereby encouraging collaboration among stakeholders.

- Ensuring Performance and Scalability: The design of the architecture allows for handling high volumes of real-time data, as generated within Industry 4.0 systems, thus providing a scalable solution. The capacity to process and manage data from multiple plants at the business DLT layer further emphasizes the architecture's scalability.

- Supporting Automation: By facilitating the implementation of smart contracts at the business DLT layer, the architecture supports the automation of business processes, such as supply chain management, product tracking, quality control, and regulatory compliance. This automation not only enhances overall operational efficiency but also reduces costs and strengthens competitiveness.

- Environmentally Conscious and Energy Efficient: The architecture's design is rooted in the intent of providing a solution that is mindful of its environmental impact and energy efficiency. By leveraging novel DLT technologies and private network solutions, which have been recognized for their potential in offering energy-efficient alternatives to traditional systems, this thesis ensured that the proposed architecture aligns with global sustainability goals.

- Cost efficient: The architecture integrates DLT to minimize operational and transactional costs. It removes the need for intermediaries, speeds up transaction processing, lowers maintenance expenses, and allows for cost-effective scalability.

In conclusion, the proposed multi-layer DLT architecture is a comprehensive solution that extends from data generation at the IoT level to its processing and utilization for business purposes at higher levels. It encapsulates the thesis' central objective, addressing the critical challenges of cybersecurity, data standardization, system interoperability, scalability, and automation within the Industry 4.0 ecosystem. Through its design, it seeks to create a secure, efficient, and transparent digital infrastructure, preserving data integrity throughout its lifecycle, fostering trust, and paving the way for advanced Industry 4.0 implementations.

5.2 List of Contributions

As a summary, the complete list of the contributions that have been made in this thesis is as follows:

- The Industry 4.0 environment was analyzed starting from the pyramid of automation standard, and a new scenario model was defined based on the aforementioned standard in order to overcome the present limitations that legacy models have towards implementing DLTs in Industry 4.0. Specifically, four main levels of industrial scenarios were defined: the machine level, the production line level, the plant level, and the consortium level. Each level encompasses several artifacts from the underlying level (e.g., a production line includes several machines). The defined scenario provides a solid base and starting point towards the design of a holistic DLT architecture for Industry 4.0.

- The state-of-the-art in DLTs applied in Industry 4.0 and other fields was analyzed. Crucial aspects of existing works, such as the technical characteristics of present architectures, open challenges, and future research opportunities were studied and discussed.

- A multi-layer DLT architecture that fulfills the main objectives of this thesis was proposed. The architecture is composed of three layers that are adjusted to the aforementioned Industry 4.0 scenario and state of the art.

  - In the first layer, called the "Data Source DLT Layer", an efficient DAG DLT was employed. Several improvements over the existing state-of-the-art DAG type DLTs were proposed. Proposed improvements affect storage, cryptography, and consensus. For storage, an IPFS decentralized storage solution was implemented to reduce the storage burden of the DLT. The performance of the cryptographic algorithms within the DLT was improved by implementing newer and more efficient algorithms. Finally, a novel reputation-based mechanism was designed to reduce the amount of PoW that lightweight devices must perform before sending transactions to the DAG DLT. With these proposed improvements, the highest possible participation rate of IIoT devices in the DLT network was aimed for. Ideally, all devices should participate as ordinary nodes to avoid fragmentation and fully take advantage of the benefits of DLTs. The performed tests and simulations demonstrated the effectiveness of the proposed improvements.
  - In the second layer, called the "Bridge DLT Layer", the challenges that arise

in a complex industrial environment at a plant level were addressed. Specifically, a "data homogenization" process was designed for solving data interoperability issues that relies on the use of decentralized blockchain oracles as a trustworthy source for the target data model scheme the data needs to conform to. A highly versatile blockchain oracle platform was employed to provide simplicity and interoperability capabilities. The resulting homogenized data was stored in a blockchain-based solution for trustworthy access and processing. A monitoring system was designed for the proposed scheme to track the quality of the retrieved data, the performance of the network, the usage of each oracle, billing reports, security incidents, etc. A monitoring architecture API for data retrieval was implemented, and it was visualized using the Elasticsearch stack. Finally, a prototype was implemented that performs the secure data homogenization process that: (i) accesses raw machine data stored in DAG DLTs, (ii) gets the target data model schema from the oracles, (iii) performs the data homogenization from the source data scheme to the target data schema, and (iv) stores the homogenized data into an interoperable "plant level" blockchain network so that it can be consistently accessed and processed by other services. The monitoring system of the aforementioned scheme was also implemented.

  - In the third layer of the architecture, called the "Business DLT Layer", an interoperable, customizable, and fee-less smart contract platform was designed. This contribution addresses the challenges of interoperability, scalability, and cost. The design allows for adaptable interfaces to accommodate various business processes, facilitating collaboration among diverse enterprises. Scalability issues are addressed by the customizability of the platform, enabling it to handle high-transaction situations without impairing performance. The challenge of cost is targeted by ensuring the platform does not require transaction fees, which can often present a substantial burden to businesses. Private channels were implemented to counter the challenge of limited data privacy. As smart contracts execute on distributed ledgers, all data used becomes transparent and accessible, which can be problematic for sensitive information. To address the challenge of limited access to external data, oracles were implemented, enabling the smart contracts to access data outside the ledger. This feature expanded the operational capability of the smart contracts, ensuring their successful execution.

- Finally, the last contribution is the implementation of the proposed architecture in a realistic use case. Specifically, the proposed solution is translated into practice

by developing the platform and demonstrating its functionality in a real-world Industry 4.0 use case developed in collaboration with a leading industrial company: Fagor Automation. This allowed to better understand how the designed platform can meet the challenges and needs of real Industry 4.0 enterprises.

5.3 List of Publications

International Journals

- **Stefanescu, D.**, Montalvillo, L., Galán-García, P., Unzilla, J., and Urbieta, A. (2022). A Systematic Literature Review of Lightweight Blockchain for IoT. IEEE Access.

- **Stefanescu, D.**, Galán-García, P., Montalvillo, L., Unzilla, J., and Urbieta, A. (2023). Industrial Data Homogenization and Monitoring Scheme with Blockchain Oracles. Smart Cities, 6(1), 263-290.

Conference Proceedings

- **Stefanescu, D.**, Galán-García, P., Montalvillo, L., Unzilla, J., and Urbieta, A. (2021, September). Towards a holistic DLT architecture for IIoT: improved DAG for production lines. In Blockchain and Applications: 3rd International Congress (pp. 179-188). Cham: Springer International Publishing.

- **Stefanescu, D.** (2022). Towards a Holistic DLT Architecture for IIoT. In Blockchain and Applications: 3rd International Congress (pp. 363-366). Springer International Publishing.

- **Stefanescu, D.**, Montalvillo, L., Galán-García, P., Unzilla, J., and Urbieta, A. (2022, July). Interoperable Industry 4.0 Plant Blockchain and Data Homogenization via Decentralized Oracles. In International Congress on Blockchain and Applications (pp. 303-313). Cham: Springer International Publishing.

- **Stefanescu, D.**, Montalvillo, L., Galán-García, P., Unzilla, J., and Urbieta, A. (2023, July). Industry 4.0 Business-oriented Blockchain Design Decision Tree. In International Congress on Blockchain and Applications. Cham: Springer International Publishing.

Sent - Under Review

- **Stefanescu, D.**, Galán-García, P., Montalvillo, L., Unzilla, J., Urbieta, A. and Caminos, J. Smart Contract Powered Framework for the Next Generation Industry 4.0 Business Model.

- Holistic DLT Architecture software registration (Spanish State)

5.4 Future Research Lines

Based on the results and findings of this thesis, the following future research lines are suggested:

Further exploration of layered DLT architectures

The proposed multi-layered DLT architecture provides a robust framework for Industry 4.0 applications, but the potential of this architectural style in other contexts remains largely unexplored. Future research can delve into implementing this type of DLT architecture for different sectors such as healthcare, logistics, smart cities, and more. Detailed case studies could be conducted to evaluate the benefits, challenges, and adaptations needed for successful implementation in these different settings.

Integration with current industrial systems and protocols

The integration of the proposed architecture with existing systems and protocols within an industrial plant is a significant area of future research. Considering the legacy systems already in use in many manufacturing industries, the proposed architecture's compatibility with such systems is vital. Thus, future research should focus on developing mechanisms for seamless integration with these systems while minimizing potential disruptions.

Exploring advanced mechanisms for data standardization and interoperability

While the proposed architecture emphasizes data standardization and system interoperability, there is always room for further exploration. Future research could focus on the development of more sophisticated methods for data standardization and the creation of more advanced interoperability protocols, potentially leveraging AI and machine learning techniques.

Extending Smart Contract capabilities

The proposed architecture suggests the use of smart contracts, primarily at the business layer. However, extending these capabilities to other layers of the architecture could open up new opportunities for automation and data management. Future work should explore the potential for the implementation of smart contracts within the data source and bridge layers, considering the specific requirements and constraints of these layers.

Quantum threats to DLTs

With the rise of quantum computing, potential threats to existing cryptographic systems have emerged. Although this thesis mentioned the exploration of post-quantum cryptographic solutions, dedicated research into quantum-resistant algorithms and their integration into DLTs for Industry 4.0 would ensure the long-term security and integrity of the system.

AI

Integrating AI within the proposed DLT architecture could further enhance its capabilities. AI can be employed to conduct sophisticated data analysis on stored data, leading to improved decision-making and process optimization. Machine Learning and Deep Learning algorithms could identify complex patterns, predicting machine failures or operational inefficiencies, providing preemptive solutions.

Lastly, AI integration could add autonomy and learning capabilities to the smart contracts, making them more adaptable and responsive. It can help in crafting sophisticated contracts that adjust based on changes in market conditions, supply chain disruptions, or changes in regulatory policies.

Regulatory and ethical implications of DLTs in Industry 4.0

Finally, as DLTs become increasingly integral to industrial systems, the regulatory landscape and ethical considerations surrounding their use will come into sharper focus. Further research into these areas, including the rights and responsibilities of different stakeholders, transparency, accountability, and data governance, would provide valuable guidance for organizations and policymakers alike.

REFERENCES

[1] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, "Industry 4.0," *Business and information systems engineering*, vol. 6, pp. 239–242, 2014.

[2] M. Kerin and D. T. Pham, "A review of emerging industry 4.0 technologies in remanufacturing," *Journal of cleaner production*, vol. 237, p. 117805, 2019.

[3] J. Schlechtendahl, M. Keinert, F. Kretschmer, A. Lechler, and A. Verl, "Making existing production systems Industry 4.0-ready: Holistic approach to the integration of existing production systems in Industry 4.0 environments," *Production Engineering*, vol. 9, no. 1, pp. 143–148, 2015.

[4] M. Rahman, A. D. Fentaye, V. Zaccaria, I. Aslanidou, E. Dahlquist, and K. Kyprianidis, "A framework for learning system for complex industrial processes," in *AI and Learning Systems-Industrial Applications and Future Directions*, IntechOpen, 2021.

[5] S. Kumar and R. R. Mallipeddi, "Impact of cybersecurity on operations and supply chain management: Emerging trends and future research directions," *Production and Operations Management*, vol. 31, no. 12, pp. 4488–4500, 2022.

[6] M. Mohamed, "Challenges and benefits of industry 4.0: An overview," *International Journal of Supply and Operations Management*, vol. 5, no. 3, pp. 256–265, 2018.

[7] B. P. Santos, A. Alberto, T. D. F. M. Lima, and F. M. B. Charrua-Santos, "Industry 4.0: challenges and opportunities," *Revista Produção e desenvolvimento*, 2018.

[8] G. Culot, F. Fattori, M. Podrecca, and M. Sartor, "Addressing industry 4.0 cybersecurity challenges," *IEEE Engineering Management Review*, vol. 47, no. 3, pp. 79–86, 2019.

[9] M. B. Mollah, J. Zhao, D. Niyato, K.-Y. Lam, X. Zhang, A. M. Ghias, L. H. Koh, and L. Yang, "Blockchain for future smart grid: A comprehensive survey," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 18–43, 2020.

[10] J. Lian, S. Wang, and Y. Xie, "Tdrb: An efficient tamper-proof detection middleware for relational database based on blockchain technology," *IEEE Access*, vol. 9, pp. 66707–66722, 2021.

[11] S. Dramé-Maigné, M. Laurent, L. Castillo, and H. Ganem, "Centralized, distributed, and everything in between: Reviewing access control solutions for the iot," *ACM Computing Surveys (CSUR)*, vol. 54, no. 7, pp. 1–34, 2021.

[12] H. Hussain, S. U. R. Malik, A. Hameed, S. U. Khan, G. Bickler, N. Min-Allah, M. B. Qureshi, L. Zhang, W. Yongji, N. Ghani, *et al.*, "A survey on resource allocation in high performance distributed computing systems," *Parallel Computing*, vol. 39, no. 11, pp. 709–736, 2013.

[13] Z. Rahman, X. Yi, S. T. Mehedi, R. Islam, and A. Kelarev, "Blockchain applicability for the internet of things: Performance and scalability challenges and solutions," *Electronics*, vol. 11, no. 9, p. 1416, 2022.

[14] K. Zhou, T. Liu, and L. Zhou, "Industry 4.0: Towards future industrial opportunities and challenges," in *2015 12th International conference on fuzzy systems and knowledge discovery (FSKD)*, pp. 2147–2152, IEEE, 2015.

[15] M. Khan, X. Wu, X. Xu, and W. Dou, "Big data challenges and opportunities in the hype of industry 4.0," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2017.

[16] T. Masood and P. Sonntag, "Industry 4.0: Adoption challenges and benefits for smes," *Computers in Industry*, vol. 121, p. 103261, 2020.

[17] M. Opazo-Basáez, F. Vendrell-Herrero, O. F. Bustinza, and J. Marić, "Global value chain breadth and firm productivity: the enhancing effect of industry 4.0," *Journal of Manufacturing Technology Management*, vol. 33, no. 4, pp. 785–804, 2022.

[18] S. Karadayi-Usta, "An interpretive structural analysis for industry 4.0 adoption challenges," *IEEE Transactions on Engineering Management*, vol. 67, no. 3, pp. 973–978, 2019.

[19] S. Hasan, C. Jatiningrum, and M. Gumanti, "The contribution of customer satisfaction towards company image and its impact in revolution industry 4.0 era," *IJEBD (International Journal of Entrepreneurship and Business Development)*, vol. 5, no. 2, pp. 359–368, 2022.

[20] K. Wüst and A. Gervais, "Do you need a blockchain?," in *2018 crypto valley conference on blockchain technology (CVCBT)*, pp. 45–54, IEEE, 2018.

[21] T. Varshney, N. Sharma, I. Kaushik, and B. Bhushan, "Authentication and encryption based security services in blockchain technology," in *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, pp. 63–68, IEEE, 2019.

[22] T. Alladi, V. Chamola, R. M. Parizi, and K. K. R. Choo, "Blockchain Applications for Industry 4.0 and Industrial IoT: A Review," *IEEE Access*, vol. 7, pp. 176935–176951, 2019.

[23] Y. Liu, K. Wang, Y. Lin, and W. Xu, "Lightchain: A lightweight blockchain system for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3571–3581, 2019.

[24] T. M. Fernandez-Carames and P. Fraga-Lamas, "A Review on the Application of Blockchain to the Next Generation of Cybersecure Industry 4.0 Smart Factories," *IEEE Access*, vol. 7, pp. 45201–45218, 2019.

[25] R. Stephen and A. Alex, "A review on blockchain security," in *IOP Conference Series: Materials Science and Engineering*, vol. 396, p. 012030, IOP Publishing, 2018.

[26] M. Isaja and J. Soldatos, "Distributed ledger technology for decentralization of manufacturing processes," in *2018 IEEE Industrial Cyber-Physical Systems (ICPS)*, pp. 696–701, IEEE, 2018.

[27] I. A. Omar, R. Jayaraman, K. Salah, M. C. E. Simsekler, I. Yaqoob, and S. Ellahham, "Ensuring protocol compliance and data transparency in clinical trials using blockchain smart contracts," *BMC Medical Research Methodology*, vol. 20, no. 1, pp. 1–17, 2020.

[28] R. Belchior, L. Riley, T. Hardjono, A. Vasconcelos, and M. Correia, "Do you need a distributed ledger technology interoperability solution?," *Distributed Ledger Technologies: Research and Practice*, vol. 2, no. 1, pp. 1–37, 2023.

[29] N. Mohamed and J. Al-Jaroodi, "Applying blockchain in industry 4.0 applications," in *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)*, pp. 0852–0858, IEEE, 2019.

[30] D. Khan, L. T. Jung, and M. A. Hashmani, "Systematic literature review of challenges in blockchain scalability," *Applied Sciences*, vol. 11, no. 20, p. 9372, 2021.

[31] C. Bai, P. Dallasega, G. Orzes, and J. Sarkis, "Industry 4.0 technologies assessment: A sustainability perspective," *International journal of production economics*, vol. 229, p. 107776, 2020.

[32] G. Llambias, B. Bradach, J. Nogueira, L. González, and R. Ruggia, "Gateway-based interoperability for dlt," *TechRxiv*, 2023.

[33] G. R. Gray and G. R. Gray, "Dlt standards," *Blockchain Technology for Managers*, pp. 165–172, 2021.

[34] G. Caldarelli, "Understanding the blockchain oracle problem: A call for action," *Information*, vol. 11, no. 11, p. 509, 2020.

[35] S. Paavolainen and P. Nikander, "Security and privacy challenges and potential solutions for dlt based iot systems," in *2018 Global Internet of Things Summit (GIoTS)*, pp. 1–6, IEEE, 2018.

[36] P. Sharma, R. Jindal, and M. D. Borah, "Blockchain technology for cloud storage: A systematic literature review," *ACM Computing Surveys (CSUR)*, vol. 53, no. 4, pp. 1–32, 2020.

[37] Y. Liu, Z. Fang, M. H. Cheung, W. Cai, and J. Huang, "Economics of blockchain storage," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2020.

[38] D. Roeck, F. Schöneseiffen, M. Greger, and E. Hofmann, "Analyzing the potential of dlt-based applications in smart factories," *Blockchain and Distributed Ledger Technology Use Cases: Applications and Lessons Learned*, pp. 245–266, 2020.

[39] R. J. Wieringa and R. J. Wieringa, "The design cycle," *Design Science Methodology for Information Systems and Software Engineering*, pp. 27–34, 2014.

[40] A. Dresch, D. P. Lacerda, J. A. V. Antunes Jr, A. Dresch, D. P. Lacerda, and J. A. V. Antunes, *Design science research*. Springer, 2015.

[41] A. R. Hevner, "A three cycle view of design science research," *Scandinavian journal of information systems*, vol. 19, no. 2, p. 4, 2007.

[42] N. Tuptuk and S. Hailes, "Security of smart manufacturing systems," *Journal of Manufacturing Systems*, vol. 47, no. February, pp. 93–106, 2018.

[43] E. M. Martinez, P. Ponce, I. Macias, and A. Molina, "Automation pyramid as constructor for a complete digital twin, case study: A didactic manufacturing system," *Sensors*, vol. 21, no. 14, p. 4656, 2021.

[44] T. P. Raptis, A. Passarella, and M. Conti, "Data management in industry 4.0: State of the art and open challenges," *IEEE Access*, vol. 7, pp. 97052–97093, 2019.

[45] U. Tariq, A. Ibrahim, T. Ahmad, Y. Bouteraa, and A. Elmogy, "Blockchain in internet-of-things: a necessity framework for security, reliability, transparency, immutability and liability," *IET Communications*, vol. 13, no. 19, pp. 3187–3192, 2019.

[46] M. Sanchez, E. Exposito, and J. Aguilar, "Industry 4.0: survey from a system integration perspective," *International Journal of Computer Integrated Manufacturing*, vol. 33, no. 10-11, pp. 1017–1041, 2020.

[47] A. A. Hijazi, S. Perera, R. N. Calheiros, and A. Alashwal, "A data model for integrating bim and blockchain to enable a single source of truth for the construction supply chain data delivery," *Engineering, Construction and Architectural Management*, 2022.

[48] C. J. Bartodziej and C. J. Bartodziej, *The concept industry 4.0.* Springer, 2017.

[49] Y. Yu, R. Chen, H. Li, Y. Li, and A. Tian, "Toward Data Security in Edge Intelligent IIoT," *IEEE Network*, vol. 33, no. 5, pp. 20–26, 2019.

[50] N. Cam-Winget, A. R. Sadeghi, and Y. Jin, "Invited - Can IoT be secured: Emerging challenges in connecting the unconnected," *Proceedings - Design Automation Conference*, vol. 05-09-June, 2016.

[51] ENISA, "Good Practices for Security of Internet of Things in the context of Smart Manufacturing," Tech. Rep. November, European Union Agency For Network and Information Security, 2018.

[52] M. e. a. Rüßmann, "Future of Productivity and Growth in Manufacturing," *Boston Consulting*, vol. 6, no. April, pp. 239–242, 2015.

[53] H. Cheng, P. Zeng, L. Xue, Z. Shi, P. Wang, and H. Yu, "Manufacturing ontology development based on industry 4.0 demonstration production line," *Proceedings - 2016 3rd International Conference on Trustworthy Systems and Their Applications, TSA 2016*, pp. 42–47, 2016.

[54] S. Weyer, M. Schmitt, M. Ohmer, and D. Gorecky, "Towards industry 4.0 - Standardization as the crucial challenge for highly modular, multi-vendor production systems," *IFAC-PapersOnLine*, vol. 28, no. 3, pp. 579–584, 2015.

[55] Z. Shi, Y. Xie, W. Xue, Y. Chen, L. Fu, and X. Xu, "Smart factory in Industry 4.0," *Systems Research and Behavioral Science*, vol. 37, no. 4, pp. 607–617, 2020.

[56] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial internet of things: Challenges, opportunities, and directions," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4724–4734, 2018.

[57] T. Lennvall, M. Gidlund, and J. Akerberg, "Challenges when bringing IoT into industrial automation," *2017 IEEE AFRICON: Science, Technology and Innovation for Africa, AFRICON 2017*, pp. 905–910, 2017.

[58] T. Stock and G. Seliger, "Opportunities of Sustainable Manufacturing in Industry 4.0," *Procedia CIRP*, vol. 40, no. Icc, pp. 536–541, 2016.

[59] X. L. Liu, W. M. Wang, H. Guo, A. V. Barenji, Z. Li, and G. Q. Huang, "Industrial blockchain based framework for product lifecycle management in industry 4.0," *Robotics and Computer-Integrated Manufacturing*, vol. 63, no. January 2019, p. 101897, 2020.

[60] A. Dorri and R. Jurdak, "Tree-Chain: A Fast Lightweight Consensus Algorithm for IoT Applications," in *2020 IEEE 45th Conference on Local Computer Networks (LCN)*, IEEE, 2020.

[61] A. Panarello, N. Tapas, G. Merlino, F. Longo, and A. Puliafito, "Blockchain and iot integration: A systematic survey," *Sensors (Switzerland)*, vol. 18, no. 8, 2018.

[62] T. M. Fernández-Caramés and P. Fraga-Lamas, "A Review on the Use of Blockchain for the Internet of Things," *IEEE Access*, vol. 6, pp. 32979–33001, 2018.

[63] M. Wu, K. Wang, X. Cai, S. Guo, M. Guo, and C. Rong, "A Comprehensive Survey of Blockchain: From Theory to IoT Applications and beyond," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8114–8154, 2019.

[64] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli, and M. H. Rehmani, "Applications of Blockchains in the Internet of Things: A Comprehensive Survey," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 2, pp. 1676–1717, 2019.

[65] M. A. Khan and K. Salah, "IoT security: Review, blockchain solutions, and open challenges," *Future Generation Computer Systems*, vol. 82, no. November, pp. 395–411, 2018.

[66] D. A. Noby and A. Khattab, "A survey of blockchain applications in IoT systems," *Proceedings - ICCES 2019: 2019 14th International Conference on Computer Engineering and Systems*, pp. 83–87, 2019.

[67] F. A. Abadi, J. Ellul, and G. Azzopardi, "The Blockchain of Things, Beyond Bitcoin: A Systematic Review," *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, no. July, pp. 1349–1354, 2018.

[68] Y. Mezquita, R. Casado, A. Gonzalez-Briones, J. Prieto, and J. M. Corchado, "Blockchain technology in IoT systems: Review of the challenges," *Annals of Emerging Technologies in Computing*, vol. 3, no. 5 Special Issue, pp. 17–24, 2019.

[69] H. N. Dai, Z. Zheng, and Y. Zhang, "Blockchain for Internet of Things: A Survey," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8076–8094, 2019.

[70] R. A. Memon, J. P. Li, J. Ahmed, M. I. Nazeer, M. Ismail, and K. Ali, "Cloud-based vs. blockchain-based IoT: a comparative survey and way forward," *Frontiers of Information Technology and Electronic Engineering*, vol. 21, no. 4, pp. 563–586, 2020.

[71] M. Conoscenti, A. Vetro, and J. C. De Martin, "Blockchain for the Internet of Things: A systematic literature review," *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*, vol. 0, no. January 2020, 2016.

[72] S. K. Lo, Y. Liu, S. Y. Chia, X. Xu, Q. Lu, L. Zhu, and H. Ning, "Analysis of Blockchain Solutions for IoT: A Systematic Literature Review," *IEEE Access*, vol. 7, pp. 58822–58835, 2019.

[73] Q. Wang, X. Zhu, Y. Ni, L. Gu, and H. Zhu, "Blockchain for the IoT and industrial IoT: A review," *Internet of Things*, vol. 10, no. xxxx, p. 100081, 2020.

[74] L. Lao, Z. Li, S. Hou, B. Xiao, S. Guo, and Y. Yang, "A survey of IoT applications in blockchain systems: Architecture, consensus, and traffic modeling," *ACM Computing Surveys*, vol. 53, no. 1, 2020.

[75] B. Farahani, F. Firouzi, and M. Luecking, "The convergence of IoT and distributed ledger technologies (DLT): Opportunities, challenges, and solutions," *Journal of Network and Computer Applications*, vol. 177, no. November 2020, p. 102936, 2021.

[76] X. Wang, X. Zha, W. Ni, R. P. Liu, Y. J. Guo, X. Niu, and K. Zheng, "Survey on blockchain for Internet of Things," *Computer Communications*, vol. 136, no. September, pp. 10–29, 2019.

[77] P. Karthikeyyan, S. Velliangiri, and I. T. Joseph, "Review of Blockchain based IoT application and its security issues," *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies, ICICICT 2019*, pp. 6–11, 2019.

[78] J. Sengupta, S. Ruj, and S. Das Bit, "A Comprehensive Survey on Attacks, Security Issues and Blockchain Solutions for IoT and IIoT," *Journal of Network and Computer Applications*, vol. 149, no. November, p. 102481, 2020.

[79] M. Alamri, N. Z. Jhanjhi, and M. Humayun, "Blockchain for Internet of Things (IoT) Research Issues Challenges and Future Directions: A Review," *International Journal of Computer Science and Network Security*, vol. 19, no. 5, pp. 244–258, 2019.

[80] M. A. Ferrag, M. Derdour, M. Mukherjee, A. Derhab, L. Maglaras, and H. Janicke, "Blockchain technologies for the internet of things: Research issues and challenges," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2188–2204, 2019.

[81] M. S. Madumidha, "Blockchain Security for Internet of Things : A Literature Survey The Internet of Things (IoT) is experiencing a tremendous growth in areas of research and industry ; however , still suffers from security issues . Conventional security mechanisms haven '," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 16, pp. 3677–3686, 2018.

[82] F. Lin, H. Hui, X. An, H. Wang, W. Ju, H. Yang, and H. Gao, "Survey on blockchain for internet of things," *Journal of Internet Services and Information Security*, vol. 9, no. 2, pp. 1–30, 2019.

[83] M. Alizadeh, K. Andersson, and O. Schelen, "A survey of secure internet of things in relation to blockchain," *Journal of Internet Services and Information Security*, vol. 10, no. 3, pp. 47–75, 2020.

[84] D. Hanggoro and R. F. Sari, "A review of lightweight blockchain technology implementation to the internet of things," *IEEE Region 10 Humanitarian Technology Conference, R10-HTC*, vol. 2019-Novem, pp. 275–280, 2019.

[85] M. A. Khan and K. Salah, "IoT security: Review, blockchain solutions, and open challenges," *Future Generation Computer Systems*, vol. 82, no. November, pp. 395–411, 2018.

[86] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, D. Altman, G. Antes, D. Atkins, V. Barbour, N. Barrowman, J. A. Berlin, J. Clark, M. Clarke, D. Cook, R. D'Amico, J. J. Deeks, P. J. Devereaux, K. Dickersin, M. Egger, E. Ernst, P. C. Gøtzsche, J. Grimshaw, G. Guyatt, J. Higgins, J. P. Ioannidis, J. Kleijnen, T. Lang, N. Magrini, D. McNamee, L. Moja, C. Mulrow, M. Napoli, A. Oxman, B. Pham, D. Rennie, M. Sampson, K. F. Schulz, P. G. Shekelle, D. Tovey, and P. Tugwell, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Medicine*, vol. 6, no. 7, 2009.

[87] K. B. and C. S., "Guidelines for performing Systematic Literature Reviews in Software Engineering," tech. rep., School of Computer Science and Mathematics, Keele University, 2007.

[88] D. M. Sheeba and S. Jayalakshmi, "Lightweight Blockchain to Improve Security and Privacy in Smarthome," *International Journal of Recent Technology and Engineering*, vol. 8, no. 6, pp. 5021–5027, 2020.

[89] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, "LSB: A Lightweight Scalable Blockchain for IoT security and anonymity," *Journal of Parallel and Distributed Computing*, vol. 134, pp. 180–197, 2019.

[90] S. N. Mohanty, K. C. Ramya, S. S. Rani, D. Gupta, K. Shankar, S. K. Lakshmanaprabu, and A. Khanna, "An efficient Lightweight integrated Blockchain (ELIB) model for IoT security and privacy," *Future Generation Computer Systems*, vol. 102, pp. 1027–1037, 2020.

[91] M. Singh, G. S. Aujla, and R. S. Bali, "ODOB: One drone one block-based lightweight blockchain architecture for internet of drones," *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS 2020*, pp. 249–254, 2020.

[92] A. Islam, T. Rahim, M. Masuduzzaman, and S. Y. Shin, "A blockchain-based artificial intelligence-empowered contagious pandemic situation supervision scheme using internet of drone things," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 166–173, 2021.

[93] A. Islam, A. Al Amin, and S. Y. Shin, "Fbi: A federated learning-based blockchain-embedded data accumulation scheme using drones for internet of things," *IEEE Wireless Communications Letters*, vol. 11, no. 5, pp. 972–976, 2022.

[94] M. A. Uddin, A. Stranieri, I. Gondal, and V. Balasurbramanian, "A lightweight blockchain based framework for underwater iot," *Electronics (Switzerland)*, vol. 8, no. 12, 2019.

[95] C. Liu, Y. Xiao, V. Javangula, Q. Hu, S. Wang, and X. Cheng, "NormaChain: A blockchain-based normalized autonomous transaction settlement system for IoT-based e-commerce," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4680–4693, 2019.

[96] J. Wang, Y. Liu, S. Niu, and H. Song, "Lightweight blockchain assisted secure routing of swarm UAS networking," *Computer Communications*, vol. 165, no. October 2020, pp. 131–140, 2021.

[97] J. Xi, S. Zou, G. Xu, and Y. Lu, "CrowdLBM: A lightweight blockchain-based model for mobile crowdsensing in the Internet of Things," *Pervasive and Mobile Computing*, vol. 84, p. 101623, 2022.

[98] L. Ismail, H. Materwala, and S. Zeadally, "Lightweight Blockchain for Healthcare," *IEEE Access*, vol. 7, pp. 149935–149951, 2019.

[99] X. Xu, Z. Zeng, S. Yang, and H. Shao, "A novel blockchain framework for industrial IoT edge computing," *Sensors (Switzerland)*, vol. 20, no. 7, pp. 1–16, 2020.

[100] T. Kim, J. Noh, and S. Cho, "SCC: Storage Compression Consensus for Blockchain in Lightweight IoT Network," *2019 IEEE International Conference on Consumer Electronics, ICCE 2019*, pp. 1–4, 2019.

[101] C. W. Huang and Y. C. Chen, "Zerocalo - A lightweight blockchain based on DHT network," *ACM International Conference Proceeding Series*, pp. 38–42, 2019.

[102] F. H. Pohrmen and G. Saha, *LightBC: A Lightweight Hash-Based Blockchain for the Secured Internet of Things*, vol. 1165. Springer Singapore, 2021.

[103] Y. Yu, S. Zhang, C. Chen, and X. Zhong, "LVChain: A lightweight and vote-based blockchain for access control in the IoT," *2018 IEEE 4th International Conference on Computer and Communications, ICCC 2018*, pp. 870–874, 2018.

[104] M. S. Siddiqui, T. A. Syed, A. Nadeem, W. Nawaz, and S. S. Albouq, "BlockTrack-L: A lightweight blockchain-based provenance message tracking in IoT," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 4, pp. 463–470, 2020.

[105] C. Connelly, S. Zaidi, M. Shakir, and H. Ahmadi, "A Lightweight Permission-Based Blockchain for IoT Environments," *2020 International Conference on UK-China Emerging Technologies (UCET)*, pp. 1–4, 2020.

[106] S. Lee, J. Lee, S. Hong, and J.-h. Kim, "Lightweight End-to-End Blockchain for IoT Applications," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 8, pp. 3224–3242, 2020.

[107] A. H. Alkhazaali and O. Ata, "Lightweight fog based solution for privacy-preserving in IoT using blockchain," *HORA 2020 - 2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*, pp. 1–10, 2020.

[108] A. Prabhakar and T. Anjali, "TCON - A lightweight Trust-dependent Consensus framework for blockchain," *2019 11th International Conference on Communication Systems and Networks, COMSNETS 2019*, vol. 2061, pp. 1–6, 2019.

[109] M. A. Y. Saputro and R. F. Sari, "Securing iot network using lightweight multi-fog (LMF) blockchain model," *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pp. 183–188, 2019.

[110] W. Yan, N. Zhang, L. L. Njilla, and X. Zhang, "PCBChain: Lightweight Reconfigurable Blockchain Primitives for Secure IoT Applications," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 28, no. 10, pp. 2196–2209, 2020.

[111] W. Zhang, Z. Wu, G. Han, Y. Feng, and L. Shu, "LDC: A lightweight dada consensus algorithm based on the blockchain for the industrial Internet of Things for smart city applications," *Future Generation Computer Systems*, vol. 108, pp. 574–582, 2020.

[112] R. Xu, Y. Chen, E. Blasch, and G. Chen, "Microchain: A Hybrid Consensus Mechanism for Lightweight Distributed Ledger for IoT," *arXiv*, pp. 1–12, 2019.

[113] Y. Xu, G. Wang, J. Yang, J. Ren, Y. Zhang, and C. Zhang, "Towards Secure Network Computing Services for Lightweight Clients Using Blockchain," *Wireless Communications and Mobile Computing*, vol. 2018, 2018.

[114] L. Xu, L. Chen, Z. Gao, S. Xu, and W. Shi, "EPBC: Efficient public blockchain client for lightweight users," *SERIAL 2017 - 1st Workshop on Scalable and Resilient Infrastructures for Distributed Ledgers, Colocated with ACM/IFIP/USENIX Middleware 2017 Conference*, pp. 1–7, 2017.

[115] R. Doku, D. B. Rawat, M. Garuba, and L. Njilla, "LightChain: On the lightweight blockchain for the internet-of-things," *Proceedings - 2019 IEEE International Conference on Smart Computing, SMARTCOMP 2019*, pp. 444–448, 2019.

[116] K. Kaur, S. Garg, G. Kaddoum, F. Gagnon, and S. H. Ahmed, "Blockchain-based lightweight authentication mechanism for vehicular fog infrastructure," *2019 IEEE International Conference on Communications Workshops, ICC Workshops 2019 - Proceedings*, no. May, pp. 20–24, 2019.

[117] S. M. Danish, M. Lestas, W. Asif, H. K. Qureshi, and M. Rajarajan, "A lightweight blockchain based two factor authentication mechanism for LoRaWAN join procedure," *2019 IEEE International Conference on Communications Workshops, ICC Workshops 2019 - Proceedings*, pp. 1–6, 2019.

[118] A. Albakri, L. Harn, and M. Maddumala, "Polynomial-based Lightweight Key Management in a Permissioned Blockchain," *2019 IEEE Conference on Communications and Network Security, CNS 2019*, pp. 1–9, 2019.

[119] Z. Yulong, N. Baoning, L. Peng, and F. Xing, *A Novel Enhanced Lightweight Node for Blockchain*, vol. 1156 CCIS. Springer, 2020.

[120] R. A. Michelin, N. Ahmed, S. S. Kanhere, A. Seneviratne, and S. Jha, "Leveraging lightweight blockchain to establish data integrity for surveillance cameras," in *2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pp. 3–5, IEEE, 2020.

[121] J. An, J. Cheng, X. Gui, W. Zhang, D. Liang, R. Gui, L. Jiang, and D. Liao, "A Lightweight Blockchain-Based Model for Data Quality Assessment in Crowdsensing," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 1, pp. 84–97, 2020.

[122] M. T. Lwin, J. Yim, and Y. B. Ko, "Blockchain-based lightweight trust management in mobile ad-hoc networks," *Sensors (Switzerland)*, vol. 20, no. 3, pp. 1–19, 2020.

[123] S. Biswas, K. Sharif, F. Li, S. Maharjan, S. P. Mohanty, and Y. Wang, "PoBT: A Lightweight Consensus Algorithm for Scalable IoT Business Blockchain," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 2343–2355, 2020.

[124] A. R. Shahid, N. Pissinou, C. Staier, and R. Kwan, "Sensor-chain: A lightweight scalable blockchain framework for internet of things," *Proceedings - 2019 IEEE International Congress on Cybermatics: 12th IEEE International Conference on Internet of Things, 15th IEEE International Conference on Green Computing and Communications, 12th IEEE International Conference on Cyber, Physical and So*, pp. 1154–1161, 2019.

[125] J. Guruprakash and S. Koppu, "EC-ElGamal and Genetic Algorithm-Based Enhancement for Lightweight Scalable Blockchain in IoT Domain," *IEEE Access*, vol. 8, pp. 141269–141281, 2020.

[126] B. Seok, J. Park, and J. H. Park, "A lightweight hash-based blockchain architecture for industrial IoT," *Applied Sciences (Switzerland)*, vol. 9, no. 18, 2019.

[127] N. H. Kim, S. M. Kang, and C. S. Hong, "Mobile charger billing system using lightweight Blockchain," *19th Asia-Pacific Network Operations and Management Symposium: Managing a World of Things, APNOMS 2017*, pp. 374–377, 2017.

[128] Y. Yu, S. Liu, P. Yeoh, B. Vucetic, and Y. Li, "LayerChain: A Hierarchical Edge-Cloud Blockchain for Large-Scale Low-Delay IIoT Applications," *IEEE Transactions on Industrial Informatics*, vol. 3203, no. c, pp. 1–1, 2020.

[129] W. Yang, X. Dai, J. Xiao, and H. Jin, "LDV: A Lightweight DAG-Based Blockchain for Vehicular Social Networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 5749–5759, 2020.

[130] Raghav, N. Andola, V. S., and V. Shekhar, "PoEWAL: A lightweight consensus mechanism for blockchain in IoT," *Elsevier Pervasive and Mobile Computing*, p. 28, 2020.

[131] W. Tiberti, A. Carmenini, L. Pomante, and D. Cassioli, "A Lightweight Blockchain-based Technique for Anti-Tampering in Wireless Sensor Networks,"

2020 23rd Euromicro Conference on Digital System Design (DSD), pp. 577–582, 2020.

[132] S. Sun, R. Du, S. Chen, and W. Li, "Blockchain-based IoT access control system: Towards security, lightweight, and cross-domain," *IEEE Access*, vol. 9, pp. 36868–36878, 2021.

[133] K. Li, Y. Yang, S. Wang, R. Shi, and J. Li, "A lightweight privacy-preserving and sharing scheme with dual-blockchain for intelligent pricing system of smart grid," *Computers and Security*, vol. 103, p. 102189, 2021.

[134] E. Bandara, D. Tosh, P. Foytik, S. Shetty, N. Ranasinghe, and K. De Zoysa, "Tikiri—Towards a lightweight blockchain for IoT," *Future Generation Computer Systems*, vol. 119, pp. 154–165, 2021.

[135] O. A. Ekanayake and M. N. Halgamuge, "Lightweight Blockchain Framework using Enhanced Master-Slave Blockchain Paradigm: Fair Rewarding Mechanism using Reward Accuracy Model," *Information Processing and Management*, vol. 58, no. 3, p. 102523, 2021.

[136] S. R. Cherupally, S. Boga, P. Podili, and K. Kataoka, "Lightweight and Scalable DAG based distributed ledger for verifying IoT data integrity," *International Conference on Information Networking*, vol. 2021-Janua, pp. 267–272, 2021.

[137] S. Khan, W. K. Lee, and S. O. Hwang, "AEchain: A lightweight blockchain for IoT applications," *IEEE Consumer Electronics Magazine*, vol. 2248, no. c, pp. 1–12, 2021.

[138] D. Na and S. Park, "Fusion chain: A decentralized lightweight blockchain for iot security and privacy," *Electronics (Switzerland)*, vol. 10, no. 4, pp. 1–18, 2021.

[139] O. Naseer, S. Ullah, and L. Anjum, "Blockchain-Based Decentralized Lightweight Control Access Scheme for Smart Grids," *Arabian Journal for Science and Engineering*, no. 0123456789, 2021.

[140] W. Lu, Z. Ren, J. Xu, and S. Chen, "Edge Blockchain Assisted Lightweight Privacy-preserving Data Aggregation for Smart Grid," *IEEE Transactions on Network and Service Management*, vol. 14, no. 8, 2020.

[141] C. Li, J. Zhang, X. Yang, and L. Youlong, "Lightweight blockchain consensus mechanism and storage optimization for resource-constrained IoT devices," *Information Processing and Management*, vol. 58, no. 4, p. 102602, 2021.

[142] S. Biswas, K. Sharif, F. Li, B. Nour, and Y. Wang, "A scalable blockchain framework for secure transactions in IoT," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4650–4659, 2019.

[143] S. Biswas, K. Sharif, F. Li, I. Alam, and S. Mohanty, "DAAC: Digital Asset Access Control in a Unified Blockchain Based E-Health System," *IEEE Transactions on Big Data*, vol. X, no. 61772077, pp. 1–15, 2020.

[144] B. Son, J. Lee, and H. Jang, "A scalable IoT protocol via an efficient dag-based distributed ledger consensus," *Sustainability (Switzerland)*, vol. 12, no. 4, pp. 1–11, 2020.

[145] L. Cui, S. Yang, Z. Chen, Y. Pan, M. Xu, and K. Xu, "An Efficient and Compacted DAG-Based Blockchain Protocol for Industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4134–4145, 2020.

[146] S. Park and H. Kim, "Dag-based distributed ledger for low-latency smart grid network," *Energies*, vol. 12, no. 18, 2019.

[147] L. Guo, J. Chen, S. Li, Y. Li, and J. Lu, "A blockchain and iot based lightweight framework for enabling information transparency in supply chain finance," *Digital Communications and Networks*, 2022.

[148] T. Baker, M. Asim, H. Samwini, N. Shamim, M. M. Alani, and R. Buyya, "A blockchain-based Fog-oriented lightweight framework for smart public vehicular transportation systems," *Computer Networks*, vol. 203, p. 108676, 2022.

[149] Q. Yao, T. Li, C. Yan, and Z. Deng, "Accident responsibility identification model for Internet of Vehicles based on lightweight blockchain," *Computational Intelligence*, no. December 2021, pp. 1–24, 2022.

[150] Z. Wang, R. Xiong, J. Jin, and C. Liang, "AirBC: A Lightweight Reputation-based Blockchain Scheme for Resource-constrained UANET," in *2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pp. 1378–1383, 2022.

[151] J. Doyle, M. Golec, and S. S. Gill, "BlockchainBus : A lightweight framework for secure virtual machine migration in cloud federations using blockchain," *Security and Privacy*, vol. 5, no. 2, pp. 1–11, 2022.

[152] J. A. Guerra, J. I. Guerrero, S. García, S. Domínguez-Cid, D. F. Larios, and C. León, "Design and Evaluation of a Heterogeneous Lightweight Blockchain-Based Marketplace," *Sensors*, vol. 22, no. 3, 2022.

[153] S. Son, J. Lee, Y. Park, Y. Park, and A. K. Das, "Design of Blockchain-Based Lightweight V2I Handover Authentication Protocol for VANET," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 3, pp. 1346–1358, 2022.

[154] Y. Jiang, X. Xu, H. Gao, A. D. Rajab, F. Xiao, and X. Wang, "LBlockchainE: A Lightweight Blockchain for Edge IoT-Enabled Maritime Transportation Systems," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–15, 2022.

[155] L. Vishwakarma, A. Nahar, and D. Das, "LBSV: Lightweight Blockchain Security Protocol for Secure Storage and Communication in SDN-Enabled IoV," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 5983–5994, 2022.

[156] B. D. Deebak, F. H. Memon, S. A. Khowaja, K. Dev, W. Wang, N. M. F. Qureshi, and C. Su, "Lightweight Blockchain Based Remote Mutual Authentication for AI-Empowered IoT Sustainable Computing Systems," *IEEE Internet of Things Journal*, p. 1, 2022.

[157] V. Mardiansyah and R. F. Sari, "Lightweight Blockchain Framework For Medical Record Data Integrity," *Journal of Applied Science and Engineering*, vol. 26, pp. 91–103, apr 2022.

[158] B. Wang and X. Hu, "Lightweight blockchain system for resource-constrained IoT devices," in *2nd International Conference on Internet of Things and Smart City (IoTSC 2022)* (F. Falcone, H. Cui, and X. Ye, eds.), vol. 12249, pp. 1–8, International Society for Optics and Photonics, SPIE, 2022.

[159] K. E. Bilami and P. LORENZ, "Lightweight Blockchain-Based Scheme to Secure Wireless M2M Area Networks," *Future Internet*, vol. 14, no. 5, 2022.

[160] M. Gupta, R. B. Patel, S. Jain, H. Garg, and B. Sharma, "Lightweight branched blockchain security framework for Internet of Vehicles," *Transactions on Emerging Telecommunications Technologies*, no. April, pp. 1–30, 2022.

[161] S. Wadhwa and Gagandeep, "Lightweight Modified Consensus Approach in IoT Blockchain," in *2022 International Conference on Emerging Smart Computing and Informatics (ESCI)*, pp. 1–5, 2022.

[162] Y. F. Ebobissé Djéné, M. S. EL Idrissi, P.-M. Tardif, B. El Bhiri, Y. Fakhri, and Y. Karfa Bekali, "Lightweight-Blockchain for Secured Wireless Sensor Networks: Energy Consumption of MAC Address-Based Proof-of-Authentication,"

in *Advanced Technologies for Humanity* (R. Saidi, B. El Bhiri, Y. Maleh, A. Mosallam, and M. Essaaidi, eds.), (Cham), pp. 182–192, Springer International Publishing, 2022.

[163] M. A. Abdullah, O. H. Alhazmi, and K. Aloufi, "Securing Internet of Things Environment using Lightweight Blockchain Approach," in *2022 4th International Conference on Applied Automation and Industrial Diagnostics (ICAAID)*, vol. 1, pp. 1–7, 2022.

[164] D. Zakariae, "A Lightweight Blockchain Framework for IoT Integration in Smart Cities," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 5, pp. 889–894, 2021.

[165] X. Qin, Y. Huang, Z. Yang, and X. Li, "LBAC: A lightweight blockchain-based access control scheme for the internet of things," *Information Sciences*, vol. 554, pp. 222–235, 2021.

[166] J.-L. Lee, P. BusiReddyGari, and B. Thompson, "A Lightweight Smart Meter Framework using a Scalable Blockchain for Smart Cities," in *2021 IEEE 7th World Forum on Internet of Things (WF-IoT)*, pp. 433–438, 2021.

[167] X. Xu and J. Peng, "A lightweight two-layer blockchain mechanism for reliable crossing-domain communication in smart cities," *arXiv preprint arXiv:2110.14860*, 2021.

[168] W. Li, M. He, W. Zhu, and J. Zheng, "A Study on Lightweight And Secure Edge Computing Based Blockchain," in *2021 IEEE 12th International Conference on Software Engineering and Service Science (ICSESS)*, pp. 256–261, 2021.

[169] Z. Wang, L. Wang, F. Xiao, Q. Chen, L. Lu, and J. Hong, "A Traditional Chinese Medicine Traceability System Based on Lightweight Blockchain," *J Med Internet Res*, vol. 23, p. e25946, jun 2021.

[170] A. A. Mamun, F. Yan, and D. Zhao, "BAASH: Lightweight, Efficient, and Reliable Blockchain-as-a-Service for HPC Systems," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '21, (New York, NY, USA), Association for Computing Machinery, 2021.

[171] H. Chai, S. Leng, J. He, K. Zhang, and B. Cheng, "CyberChain: Cybertwin Empowered Blockchain for Lightweight and Privacy-Preserving Authentication

in Internet of Vehicles," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 4620–4631, 2022.

[172] R. Zhang, M. Song, T. Li, Z. Yu, Y. Dai, X. Liu, and G. Wang, "Democratic learning: hardware/software co-design for lightweight blockchain-secured on-device machine learning," *Journal of Systems Architecture*, vol. 118, p. 102205, 2021.

[173] Q. Xie, F. Dong, and X. Feng, "ECLB: Edge-Computing-Based Lightweight Blockchain Framework for Mobile Systems," *Security and Communication Networks*, vol. 2021, p. 5510586, 2021.

[174] F. H. Pohrmen and G. Saha, "LightBC: A Lightweight Hash-Based Blockchain for the Secured Internet of Things," in *International Conference on Innovative Computing and Communications* (D. Gupta, A. Khanna, S. Bhattacharyya, A. E. Hassanien, S. Anand, and A. Jaiswal, eds.), (Singapore), pp. 811–819, Springer Singapore, 2021.

[175] J. Yuan and L. Njilla, "Lightweight and Reliable Decentralized Reward System using Blockchain," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6, 2021.

[176] N. Ding and Y. Zhao, "Lightweight Blockchain Based on Storage Resource Optimization for Internet of Vehicles," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 1063–1068, 2021.

[177] C. Li, J. Zhang, X. Yang, and L. Youlong, "Lightweight blockchain consensus mechanism and storage optimization for resource-constrained IoT devices," *Information Processing and Management*, vol. 58, no. 4, p. 102602, 2021.

[178] J. P. Mehare and M. M. Bartere, "Lightweight Blockchain Secured Framework for Smart Precise Farming System," in *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, pp. 1–6, 2021.

[179] D. Na and S. Park, "Lightweight blockchain to solve forgery and privacy issues of vehicle image data," in *2021 22nd Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 37–40, 2021.

[180] H. H. Saeed, A. B. Masood, and H. K. Qureshi, "LSM: A Lightweight Security Mechanism for IoT Based Smart City Management Systems using Blockchain," *International Journal of Innovations in Science and Technology*, vol. 3, no. 4, pp. 1–14, 2021.

[181] H. Materwala and L. Ismail, "Secure and Privacy-Preserving Lightweight Blockchain for Energy Trading," in *2021 8th International Conference on Future Internet of Things and Cloud (FiCloud)*, pp. 394–399, 2021.

[182] X. Zhang, R. Li, W. Hou, and H. Zhao, "V-Lattice: A Lightweight Blockchain Architecture Based on DAG-Lattice Structure for Vehicular Ad Hoc Networks," *Security and Communication Networks*, vol. 2021, 2021.

[183] A. D. Dwivedi, R. Singh, S. Dhall, G. Srivastava, and S. K. Pal, "Tracing the source of fake news using a scalable blockchain distributed network," in *2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pp. 38–43, 2020.

[184] B. Deebak, F. H. Memon, S. A. Khowaja, K. Dev, W. Wang, N. M. F. Qureshi, and C. Su, "A lightweight blockchain-based remote mutual authentication for ai-empowered iot sustainable computing systems," *IEEE Internet of Things Journal*, vol. 10, no. 8, pp. 6652–6660, 2022.

[185] L. Vishwakarma, A. Nahar, and D. Das, "Lbsv: Lightweight blockchain security protocol for secure storage and communication in sdn-enabled iov," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 5983–5994, 2022.

[186] M. Amiri-Zarandi, R. A. Dara, and E. Fraser, "Lbtm: A lightweight blockchain-based trust management system for social internet of things," *The Journal of Supercomputing*, pp. 1–19, 2022.

[187] Y. Jiang, X. Xu, H. Gao, A. D. Rajab, F. Xiao, and X. Wang, "Lblockchaine: A lightweight blockchain for edge iot-enabled maritime transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 2307–2321, 2022.

[188] M. Jagdish, D. U. Shah, V. Agarwal, G. B. Loganathan, A. Alqahtani, S. A. Rahin, *et al.*, "Identification of end-user economical relationship graph using lightweight blockchain-based bert model," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[189] A. El Azzaoui, M. Y. Choi, C. H. Lee, and J. H. Park, "Scalable lightweight blockchain-based authentication mechanism for secure voip communication," *Hum.-Cent. Comput. Inf. Sci*, vol. 12, no. 8, 2022.

[190] O. Said, "Lbss: A lightweight blockchain-based security scheme for iot-enabled healthcare environment," *Sensors*, vol. 22, no. 20, p. 7948, 2022.

[191] M. Zhang, R. Cao, F. Duan, Y. Yang, Y. Lu, M. Zhang, and X. Lu, "Research on lightweight blockchain technology based on edge computing," in *2022 7th IEEE International Conference on Data Science in Cyberspace (DSC)*, pp. 533–539, IEEE, 2022.

[192] V. Mardiansyah, R. F. Sari, *et al.*, "Lightweight blockchain framework for medical record data integrity," *Journal of Applied Science and Engineering*, vol. 26, no. 1, pp. 91–103, 2022.

[193] S. S. Hameedi and O. Bayat, "Improving iot data security and integrity using lightweight blockchain dynamic table," *Applied Sciences*, vol. 12, no. 18, p. 9377, 2022.

[194] K. E. Bilami and P. Lorenz, "Lightweight blockchain-based scheme to secure wireless m2m area networks," *Future Internet*, vol. 14, no. 5, p. 158, 2022.

[195] Y. F. Ebobissé Djéné, M. S. El Idrissi, P.-M. Tardif, A. Jorio, B. El Bhiri, and Y. Fakhri, "A formal energy consumption analysis to secure cluster-based wsn: A case study of multi-hop clustering algorithm based on spectral classification using lightweight blockchain," *Sensors*, vol. 22, no. 20, p. 7730, 2022.

[196] M. A. Abdullah, O. H. Alhazmi, and K. Aloufi, "Securing internet of things environment using lightweight blockchain approach," in *2022 4th International Conference on Applied Automation and Industrial Diagnostics (ICAAID)*, vol. 1, pp. 1–7, IEEE, 2022.

[197] S. Kably, M. Arioua, and N. Alaoui, "Lightweight direct acyclic graph blockchain for enhancing resource-constrained iot environment.," *Computers, Materials and Continua*, vol. 71, no. 3, 2022.

[198] S. Wang, H. Li, J. Chen, J. Wang, and Y. Deng, "Dag blockchain-based lightweight authentication and authorization scheme for iot devices," *Journal of Information Security and Applications*, vol. 66, p. 103134, 2022.

[199] J. Ktari, T. Frikha, F. Chaabane, M. Hamdi, and H. Hamam, "Agricultural lightweight embedded blockchain system: a case study in olive oil," *Electronics*, vol. 11, no. 20, p. 3394, 2022.

[200] S. Selvarajan, G. Srivastava, A. O. Khadidos, A. O. Khadidos, M. Baza, A. Alshehri, and J. C.-W. Lin, "An artificial intelligence lightweight blockchain security model for security and privacy in iiot systems," *Journal of Cloud Computing*, vol. 12, no. 1, p. 38, 2023.

[201] L. Settipalli, G. Gangadharan, and S. Bellamkonda, "An extended lightweight blockchain based collaborative healthcare system for fraud prevention," *Cluster Computing*, pp. 1–11, 2023.

[202] R. Jin, J. Hu, G. Min, and J. Mills, "Lightweight blockchain-empowered secure and efficient federated edge learning," *IEEE Transactions on Computers*, 2023.

[203] N. Yang, D. Guo, Y. Jiao, G. Ding, and T. Qu, "Lightweight blockchain-based secure spectrum sharing in space-air-ground integrated iot network," *IEEE Internet of Things Journal*, 2023.

[204] M. Maroufi, R. Abdolee, B. M. Tazekand, and S. A. Mortazavi, "Lightweight blockchain-based architecture for 5g enabled iot," *IEEE Access*, 2023.

[205] M. A. Mahmoud, M. Gurunathan, R. Ramli, K. A. Babatunde, and F. H. Faisal, "Review and development of a scalable lightweight blockchain integrated model (lightblock) for iot applications," *Electronics*, vol. 12, no. 4, p. 1025, 2023.

[206] A. A. Laghari, A. A. Khan, R. Alkanhel, H. Elmannai, and S. Bourouis, "Lightweight-biov: blockchain distributed ledger technology (bdlt) for internet of vehicles (iovs)," *Electronics*, vol. 12, no. 3, p. 677, 2023.

[207] P. Hegde and P. K. R. Maddikunta, "Secure pbft consensus-based lightweight blockchain for healthcare application," *Applied Sciences*, vol. 13, no. 6, p. 3757, 2023.

[208] R. Raj and M. Ghosh, "A lightweight blockchain framework for secure transaction in resource constrained iot devices," in *2023 5th International Conference on Recent Advances in Information Technology (RAIT)*, pp. 1–7, IEEE, 2023.

[209] O. Cheikhrouhou, K. Mershad, F. Jamil, R. Mahmud, A. Koubaa, and S. R. Moosavi, "A lightweight blockchain and fog-enabled secure remote patient monitoring system," *Internet of Things*, vol. 22, p. 100691, 2023.

[210] T. Yu, S. Yi, and L. Zhaowen, "Tinyledger: A lightweight blockchain ledger protocol for the mec network," *Computers and Electrical Engineering*, vol. 109, p. 108749, 2023.

[211] M. Divya and N. B. Biradar, "IOTA-Next Generation Block chain," *International Journal Of Engineering And Computer Science*, vol. 7, no. 04, pp. 23823–23826, 2018.

[212] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," tech. rep., Bitcoin, 2009.

[213] A. Gervais, G. O. Karame, K. Wüst, and H. Ritzdorf, "On the Security and Performance of Proof of Work Blockchains Vasileios Glykantzis Srdjaň Capkun," *Bitcoin.org*, pp. 3–16, 2017.

[214] H. Sukhwani, J. M. Martínez, X. Chang, K. S. Trivedi, and A. Rindos, "Performance modeling of PBFT consensus process for permissioned blockchain network (hyperledger fabric)," *Proceedings of the IEEE Symposium on Reliable Distributed Systems*, vol. 2017-Septe, no. March 2018, pp. 253–255, 2017.

[215] M. Raikwar, D. Gligoroski, and K. Kralevska, "SoK of Used Cryptography in Blockchain," *IEEE Access*, vol. 7, pp. 148550–148575, 2019.

[216] D. Mourtzis, E. Vlachou, and N. Milas, "Industrial Big Data as a Result of IoT Adoption in Manufacturing," *Procedia CIRP*, vol. 55, pp. 290–295, 2016.

[217] J. Truby, "Decarbonizing Bitcoin: Law and policy choices for reducing the energy consumption of Blockchain technologies and digital currencies," *Energy Research and Social Science*, vol. 44, no. February, pp. 399–410, 2018.

[218] X. Liu and N. Ansari, "Toward Green IoT: Energy Solutions and Key Challenges," *IEEE Communications Magazine*, vol. 57, no. 3, pp. 104–110, 2019.

[219] C. Cachin, "Blockchain, Cryptography, and Consensus," *Electronic Proceedings in Theoretical Computer Science*, vol. 261, no. June, pp. 1–1, 2017.

[220] M. Ali, H. Karimipour, and M. Tariq, "Integration of blockchain and federated learning for internet of things: Recent advances and future challenges," *Computers and Security*, vol. 108, p. 102355, 2021.

[221] T. M. Fernandez-Carames and P. Fraga-Lamas, "Towards Post-Quantum Blockchain: A Review on Blockchain Cryptography Resistant to Quantum Computing Attacks," *IEEE Access*, vol. 8, pp. 21091–21116, 2020.

[222] N. El Ioini and C. Pahl, "A Review of Distributed Ledger Technologies," in *On the Move to Meaningful Internet Systems. OTM 2018 Conferences* (H. Panetto, C. Debruyne, H. A. Proper, C. A. Ardagna, D. Roman, and R. Meersman, eds.), (Cham), pp. 277–288, Springer International Publishing, 2018.

[223] C. Sguanci, R. Spatafora, and A. M. Vergani, "Layer 2 Blockchain Scaling: a Survey," *CoRR*, vol. abs/2107.1, 2021.

[224] Q. Zhou, H. Huang, Z. Zheng, and J. Bian, "Solutions to Scalability of Blockchain: a Survey," *IEEE Access*, vol. 8, pp. 16440–16455, 2020.

[225] A. Kumari, S. Tanwar, S. Tyagi, and N. Kumar, "Blockchain-Based Massive Data Dissemination Handling in IIoT Environment," *IEEE Network*, vol. 35, no. 1, pp. 318–325, 2021.

[226] J. Sedlmeir, H. U. Buhl, G. Fridgen, and R. Keller, "The Energy Consumption of Blockchain Technology: Beyond Myth," *Business and Information Systems Engineering*, vol. 62, no. 6, pp. 599–608, 2020.

[227] R. Yang, R. Wakefield, S. Lyu, S. Jayasuriya, F. Han, X. Yi, X. Yang, G. Amarasinghe, and S. Chen, "Public and private blockchain in construction business process and information integration," *Automation in Construction*, vol. 118, no. May, p. 103276, 2020.

[228] S. Popov, "The Tangle," tech. rep., IOTA Foundation, 2018.

[229] W. F. Silvano and R. Marcelino, "Iota Tangle: A cryptocurrency to communicate Internet-of-Things data," *Future Generation Computer Systems*, vol. 112, pp. 307–319, 2020.

[230] M. Pincheira, M. Vecchio, R. Giaffreda, and S. S. Kanhere, "Exploiting constrained IoT devices in a trustless blockchain-based water management system," *IEEE International Conference on Blockchain and Cryptocurrency, ICBC 2020*, 2020.

[231] X. Ouyang, N. S. Islam, R. Rajachandrasekar, J. Jose, M. Luo, H. Wang, and D. K. Panda, "SSD-Assisted Hybrid Memory to Accelerate Memcached over High Performance Networks *," *41st International Conference on Parallel Processing*, pp. 470–479, 2012.

[232] M. Naz, F. A. Al-zahrani, R. Khalid, N. Javaid, A. M. Qamar, M. K. Afzal, and M. Shafiq, "A Secure Data Sharing Platform Using Blockchain and Interplanetary File System," *Sustainability (Switzerland)*, vol. 11, no. 24, 2019.

[233] J. Shen, Y. Li, Y. Zhou, and X. Wang, "Understanding I/O performance of IPFS storage: A client's perspective," *Proceedings of the International Symposium on Quality of Service*, 2019.

[234] A. H. Al-Ahdal and N. K. Deshmukh, "A systematic technical survey of lightweight cryptography on Iot environment," *International Journal of Scientific and Technology Research*, vol. 9, no. 3, pp. 6246–6261, 2020.

[235] J. P. Aumasson, L. Henzen, W. Meier, and M. Naya-Plasencia, "Quark: A lightweight hash," *Journal of Cryptology*, vol. 26, no. 2, pp. 313–339, 2013.

[236] W. Wu, S. Wu, L. Zhang, J. Zou, and L. Dong, "Spongent: A lightweight hash function," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8567, pp. 291–308, 2014.

[237] W. Fang, W. Chen, W. Zhang, J. Pei, W. Gao, and G. Wang, "Digital signature scheme for information non-repudiation in blockchain: a state of the art review," *Eurasip Journal on Wireless Communications and Networking*, vol. 2020, no. 1, 2020.

[238] D. J. Bernstein, N. Duif, T. Lange, P. Schwabe, and B. Y. Yang, "High-speed high-security signatures," *Journal of Cryptographic Engineering*, vol. 2, no. 2, pp. 77–89, 2012.

[239] NIST, "Post-Quantum Cryptography - Round 3 Submissions," 2021.

[240] D. Soni, K. Basu, M. Nabeel, N. Aaraj, M. Manzano, and R. Karri, "FALCON," in *Hardware Architectures for Post-Quantum Digital Signature Schemes*, pp. 31–41, Cham: Springer International Publishing, 2021.

[241] D. Kuemper, T. Iggena, R. Toenjes, and E. Pulvermueller, "Valid.IoT," in *MMSys '18: Proceedings of the 9th ACM Multimedia Systems Conference*, pp. 294–303, 2018.

[242] V. Jirkovsky, M. Obitko, and V. Marik, "Understanding data heterogeneity in the context of cyber-physical systems integration," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 660–667, 2017.

[243] R. Roman, P. Najera, and J. Lopez, "Securing the Internet of Things," *Computer*, vol. 44, no. 9, pp. 51–58, 2011.

[244] G. Büchi, M. Cugno, and R. Castagnoli, "Smart factory performance and industry 4.0," *Technological Forecasting and Social Change*, vol. 150, p. 119790, 2020.

[245] U. Bodkhe, S. Tanwar, K. Parekh, P. Khanpara, S. Tyagi, N. Kumar, and M. Alazab, "Blockchain for Industry 4.0: A comprehensive review," *IEEE Access*, vol. 8, pp. 79764–79800, 2020.

[246] I. Weber, X. Xu, R. Riveret, G. Governatori, A. Ponomarev, and J. Mendling, "Untrusted business process monitoring and execution using blockchain," in *Business Process Management* (M. La Rosa, P. Loos, and O. Pastor, eds.), (Cham), pp. 329–347, Springer International Publishing, 2016.

[247] D. Raposo, A. Rodrigues, S. Sinche, J. Sá Silva, and F. Boavida, "Industrial iot monitoring: Technologies and architecture proposal," *Sensors*, vol. 18, no. 10, 2018.

[248] L. Zhao, I. Brandao Machado Matsuo, Y. Zhou, and W.-J. Lee, "Design of an industrial iot-based monitoring system for power substations," *IEEE Transactions on Industry Applications*, vol. 55, no. 6, pp. 5666–5674, 2019.

[249] R. P. George, B. L. Peterson, O. Yaros, D. L. Beam, J. M. Dibbell, and R. C. Moore, "Blockchain for business," *Journal of Investment Compliance*, vol. 20, no. 1, pp. 17–21, 2019.

[250] R. Drath and A. Horch, "Industrie 4.0: Hit or hype?[industry forum]," *IEEE industrial electronics magazine*, vol. 8, no. 2, pp. 56–58, 2014.

[251] E. Oztemel and S. Gursev, "Literature review of industry 4.0 and related technologies," *Journal of intelligent manufacturing*, vol. 31, pp. 127–182, 2020.

[252] X. Xu, I. Weber, M. Staples, L. Zhu, J. Bosch, L. Bass, C. Pautasso, and P. Rimba, "A taxonomy of blockchain-based systems for architecture design," in *2017 IEEE international conference on software architecture (ICSA)*, pp. 243–252, IEEE, 2017.

[253] T. A. Almeshal and A. A. Alhogail, "Blockchain for businesses: A scoping review of suitability evaluations frameworks," *IEEE Access*, vol. 9, pp. 155425–155442, 2021.

[254] G. Prause, "Smart contracts for smart supply chains," *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 2501–2506, 2019.

[255] P. Cuccuru, "Beyond bitcoin: an early overview on smart contracts," *International Journal of Law and Information Technology*, vol. 25, no. 3, pp. 179–195, 2017.

[256] W. Zou, D. Lo, P. S. Kochhar, X.-B. D. Le, X. Xia, Y. Feng, Z. Chen, and B. Xu, "Smart contract development: Challenges and opportunities," *IEEE Transactions on Software Engineering*, vol. 47, no. 10, pp. 2084–2106, 2021.

[257] Z. Zheng, S. Xie, H.-N. Dai, W. Chen, X. Chen, J. Weng, and M. Imran, "An overview on smart contracts: Challenges, advances and platforms," *Future Generation Computer Systems*, vol. 105, pp. 475–491, 2020.

[258] M. H. Tabatabaei, R. Vitenberg, and N. R. Veeraragavan, "Understanding blockchain: definitions, architecture, design, and system comparison," *arXiv preprint arXiv:2207.02264*, 2022.

[259] D. Stefanescu, P. Galán-García, L. Montalvillo, J. Unzilla, and A. Urbieta, "Industrial data homogenization and monitoring scheme with blockchain oracles," *Smart Cities*, vol. 6, no. 1, pp. 263–290, 2023.

[260] M. Conti, G. Kumar, P. Nerurkar, R. Saha, and L. Vigneri, "A survey on security challenges and solutions in the iota," *Journal of Network and Computer Applications*, p. 103383, 2022.

[261] D. Guegan, "Public blockchain versus private blockchain," *HAL SHS*, 2017.

[262] O. Choudhury, I. Sylla, N. Fairoza, and A. Das, "A blockchain framework for ensuring data quality in multi-organizational clinical trials," in *2019 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 1–9, 2019.

[263] W. She, Z.-H. Gu, X.-K. Lyu, Q. Liu, Z. Tian, and W. Liu, "Homomorphic consortium blockchain for smart home system sensitive data privacy preserving," *IEEE Access*, vol. 7, pp. 62058–62070, 2019.

[264] V. R. Basili, R. W. Selby, and D. H. Hutchens, "Experimentation in software engineering," *Experimentation in Software Engineering*, vol. SE-12, no. 7, pp. 733–743, 1986.

[265] M. Zander, T. Waite, and D. Harz, "DAGsim," *ACM SIGMETRICS Performance Evaluation Review*, vol. 46, no. 3, pp. 118–121, 2019.

[266] O. Ascigil, S. Reñé, M. Król, G. Pavlou, L. Zhang, T. Hasegawa, Y. Koizumi, and K. Kita, "Towards peer-to-peer content retrieval markets: Enhancing IPFs with ICN," *ICN 2019 - Proceedings of the 2019 Conference on Information-Centric Networking*, pp. 78–88, 2019.

[267] M. Bhandary, M. Parmar, and D. Ambawade, "A blockchain solution based on directed acyclic graph for iot data security using iota tangle," in *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, pp. 827–832, 2020.

[268] A. F. Rochim, M. A. Aziz, and A. Fauzi, "Design log management system of computer network devices infrastructures based on elk stack," in *2019 International Conference on Electrical Engineering and Computer Science (ICECOS)*, pp. 338–342, 2019.

[269] A. Majeed, J. Lv, and T. Peng, "A framework for big data driven process analysis and optimization for additive manufacturing," *Rapid Prototyping Journal*, 2018.

[270] W. Zhao, C. Jiang, H. Gao, S. Yang, and X. Luo, "Blockchain-enabled cyber–physical systems: A review," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4023–4034, 2020.

[271] M. Bartholic, A. Laszka, G. Yamamoto, and E. W. Burger, "A taxonomy of blockchain oracles: The truth depends on the question," in *2022 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pp. 1–15, IEEE, 2022.

[272] Y. Jiang, C. Wang, Y. Wang, and L. Gao, "A cross-chain solution to integrating multiple blockchains for IoT data management," *Sensors (Switzerland)*, vol. 19, no. 9, pp. 1–18, 2019.

[273] E. Androulaki, A. Barger, V. Bortnikov, C. Cachin, K. Christidis, A. De Caro, D. Enyeart, C. Ferris, G. Laventman, Y. Manevich, *et al.*, "Hyperledger fabric: a distributed operating system for permissioned blockchains," in *Proceedings of the thirteenth EuroSys conference*, pp. 1–15, 2018.

[274] D. Stefanescu, P. Galán-García, L. Montalvillo, J. Unzilla, and A. Urbieta, "Towards a Holistic DLT Architecture for IIoT: Improved DAG for Production Lines," in *Blockchain and Applications* (J. Prieto, A. Partida, P. Leitão, and A. Pinto, eds.), (Cham), pp. 179–188, Springer International Publishing, 2022.

[275] D. Stefanescu, L. Montalvillo, P. Galán-García, J. Unzilla, and A. Urbieta, "Interoperable industry 4.0 plant blockchain and data homogenization via decentralized oracles," in *International Congress on Blockchain and Applications*, pp. 303–313, Springer, 2023.

[276] P. Belohlavek, *OEE: overall equipment effectiveness*. Blue Eagle Group, 2006.

[277] O. Dib, K.-L. Brousmiche, A. Durand, E. Thea, and E. B. Hamida, "Consortium blockchains: Overview, applications and challenges," *International Journal On Advances in Telecommunications*, vol. 11, no. 1and2, pp. 51–64, 2018.

[278] H. Boyes, B. Hallaq, J. Cunningham, and T. Watson, "The industrial internet of things (iiot): An analysis framework," *Computers in industry*, vol. 101, pp. 1–12, 2018.

[279] R. Strange and A. Zucchella, "Industry 4.0, global value chains and international business," *Multinational Business Review*, vol. 25, no. 3, pp. 174–184, 2017.

[280] I. Henao-Hernández, E. L. Solano-Charris, A. Muñoz-Villamizar, J. Santos, and R. Henríquez-Machado, "Control and monitoring for sustainable manufacturing in the industry 4.0: A literature review," *IFAC-PapersOnLine*, vol. 52, no. 10, pp. 195–200, 2019. 13th IFAC Workshop on Intelligent Manufacturing Systems IMS 2019.

[281] G.-C. Pătru, D.-C. Trancă, C.-M. Costea, D. Rosner, and R.-V. Rughiniş, "Lora based, low power remote monitoring and control solution for industry 4.0 factories and facilities," in *2019 18th RoEduNet Conference: Networking in Education and Research (RoEduNet)*, pp. 1–6, 2019.

[282] J. Gao, E. Zhu, S. Shim, and L. Chang, "Monitoring software components and component-based software," in *Proceedings 24th Annual International Computer Software and Applications Conference. COMPSAC2000*, pp. 403–412, 2000.

[283] A. van Hoorn, J. Waller, and W. Hasselbring, "Kieker: A framework for application performance monitoring and dynamic software analysis," in *Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering*, ICPE '12, (New York, NY, USA), p. 247–248, Association for Computing Machinery, 2012.

[284] M. Di Pierro, "What is the blockchain?," *Computing in Science and Engineering*, vol. 19, no. 5, pp. 92–95, 2017.

[285] S. Schulte, M. Sigwart, P. Frauenthaler, and M. Borkowski, "Towards blockchain interoperability," in *Business Process Management: Blockchain and Central and Eastern Europe Forum* (C. Di Ciccio, R. Gabryelczyk, L. García-Bañuelos, T. Hernaus, R. Hull, M. Indihar Štemberger, A. Kő, and M. Staples, eds.), (Cham), pp. 3–10, Springer International Publishing, 2019.

[286] R. Belchior, A. Vasconcelos, S. Guerreiro, and M. Correia, "A Survey on Blockchain Interoperability: Past, Present, and Future Trends," *ACM Comput. Surv.*, vol. 54, 10 2021.

[287] F. Chiacchio, D. D'Urso, L. Compagno, M. Chiarenza, and L. Velardita, "Towards a blockchain based traceability process: A case study from pharma industry," in *Advances in Production Management Systems. Production Management for the Factory of the Future* (F. Ameri, K. E. Stecke, G. von Cieminski, and D. Kiritsis, eds.), (Cham), pp. 451–457, Springer International Publishing, 2019.

[288] Z. Zheng, S. Xie, H.-N. Dai, X. Chen, and H. Wang, "Blockchain Challenges and Opportunities : A Survey Shaoan Xie Hong-Ning Dai Huaimin Wang," *International Journal of Web and Grid Services*, vol. 14, no. 4, pp. 1–24, 2017.

[289] M. Du, X. Ma, Z. Zhang, X. Wang, and Q. Chen, "A review on consensus algorithm of blockchain," *2017 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2017*, vol. 2017-Janua, pp. 2567–2572, 2017.

[290] T. Xue, Y. Yuan, Z. Ahmed, K. Moniz, G. Cao, and C. Wang, "Proof of Contribution: A Modification of Proof of Work to Increase Mining Efficiency," *Proceedings - International Computer Software and Applications Conference*, vol. 1, pp. 636–644, 2018.

[291] F. Saleh, "Blockchain without Waste: Proof-of-Stake," *The Review of Financial Studies*, vol. 34, pp. 1156–1190, 2020.

[292] M. Castro and B. Liskov, "Practical byzantine fault tolerance," *OSDI*, vol. 99, pp. 173–186, 1999.

[293] H. D. Zubaydi, Y. W. Chong, K. Ko, S. M. Hanshi, and S. Karuppayah, "A review on the role of blockchain technology in the healthcare domain," *Electronics (Switzerland)*, vol. 8, no. 6, pp. 1–29, 2019.

[294] A. Wahab and W. Memood, "Survey of consensus protocols," *arXiv*, pp. 1–12, 2018.

[295] S. D. Angelis, L. Aniello, R. Baldoni, F. Lombardi, A. Margheri, and V. Sassone, "PBFT vs proof-of-authority: applying the CAP theorem to permissioned blockchain," in *Italian Conference on Cyber Security (06/02/18)*, 2018.

[296] L. Chen, L. Xu, N. Shah, Z. Gao, Y. Lu, and W. Shi, "On security analysis of proof-of-elapsed-time (PoET)," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10616 LNCS, no. May 2019, pp. 282–297, 2017.

[297] I. Bentov, C. Lee, A. Mizrahi, and M. Rosenfeld, "Proof of Activity," *ACM SIGMETRICS Performance Evaluation Review*, vol. 42, no. 3, pp. 34–37, 2014.

[298] D. Huang, X. Ma, and S. Zhang, "Performance analysis of the raft consensus algorithm for private blockchains," *arXiv*, vol. 50, no. 1, pp. 172–181, 2018.

[299] K. Karantias, A. Kiayias, and D. Zindros, "Proof-of-Burn," in *International Conference on Financial Cryptography and Data Security*, pp. 523–540, Springer, Cham, 2020.

[300] N. Szabo, "Smart contracts: building blocks for digital markets," *EXTROPY: The Journal of Transhumanist Thought*, vol. 18, no. 2, p. 28, 1996.

[301] D. Huang, X. Ma, and S. Zhang, "Performance analysis of the raft consensus algorithm for private blockchains," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 1, pp. 172–181, 2020.

[302] M. Conner, "Sensors empower the" Internet of Things"," *EDN (Electrical Design News)*, vol. 55, no. 10, p. 32, 2010.

[303] T. Ladd and O. Groth, "Future Internet: The Internet of Things," in *3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*, vol. 5, pp. 376–380, IEEE, 2010.

[304] H. Kopetz, "Internet of Things, Real-Time Systems," *International Journal of Innovations and Advancement in Computer Science*, vol. 3, no. 8, pp. 1–20, 2011.

[305] A. R. Biswas and R. Giaffreda, "IoT and cloud convergence: Opportunities and challenges," *2014 IEEE World Forum on Internet of Things, WF-IoT 2014*, pp. 375–376, 2014.

[306] S. A. Al-Qaseemi, H. A. Almulhim, M. F. Almulhim, and S. R. Chaudhry, "IoT architecture challenges and issues: Lack of standardization," *FTC 2016 - Proceedings of Future Technologies Conference*, no. December, pp. 731–738, 2017.

[307] M. A. Uddin, A. Stranieri, I. Gondal, and V. Balasurbramanian, "A Survey on the Adoption of Blockchain in IoT: Challenges and Solutions," *Blockchain: Research and Applications*, p. 100006, 2021.

[308] H. Al-Breiki, M. H. U. Rehman, K. Salah, and D. Svetinovic, "Trustworthy Blockchain Oracles: Review, Comparison, and Open Research Challenges," *IEEE Access*, vol. 8, pp. 85675–85685, 2020.

[309] K. Mammadzada, M. Iqbal, F. Milani, L. García-Bañuelos, and R. Matulevičius, "Blockchain oracles: A framework for blockchain-based applications," in *Business Process Management: Blockchain and Robotic Process Automation Forum*

(A. Asatiani, J. M. García, N. Helander, A. Jiménez-Ramírez, A. Koschmider, J. Mendling, G. Meroni, and H. A. Reijers, eds.), (Cham), pp. 19–34, Springer International Publishing, 2020.

[310] M. D. Sheldon, "Auditing the Blockchain Oracle Problem," *Journal of Information Systems*, vol. 35, pp. 121–133, 06 2020.

[311] S. K. Ezzat, Y. N. Saleh, and A. A. Abdel-Hamid, "Blockchain oracles: State-of-the-art and research directions," *IEEE Access*, 2022.

[312] G. Caldarelli, "Overview of blockchain oracle research," *Future Internet*, vol. 14, no. 6, 2022.

[313] A. Lohachab, S. Garg, B. Kang, M. B. Amin, J. Lee, S. Chen, and X. Xu, "Towards Interconnected Blockchains: A Comprehensive Review of the Role of Interoperability among Disparate Blockchains," *ACM Comput. Surv.*, vol. 54, no. 7, 2021.

[314] G. Wood, "POLKADOT: VISION FOR A HETEROGENEOUS MULTI-CHAIN FRAMEWORK," tech. rep., Eindhoven University of Technology (TU/e), 2016.

Appendices

Appendix A

Background - Core Concepts

A.1 Industry 4.0

The Fourth Industrial Revolution, known as Industry 4.0, is the ongoing shift towards automation and data exchange across various sectors. It leverages advanced technologies such as AI, cloud computing, data analytics, and IoT to transform traditional industries [1]. A central concept within this revolution is the emergence of "smart factories", where interconnected machines form a complete visualization of the production line, making autonomous decisions. These smart factories operate through a network of devices and machinery that harness the power of the IIoT, AI, and data analytics. Machines engage in M2M communication, generating a flow of data that can prompt immediate adjustments and inform future improvements. Predictive maintenance, a notable feature, uses sensor data to preemptively service machinery, thereby reducing downtime and maximizing efficiency.

Industry 4.0 brings significant efficiency improvements and opens new avenues for businesses, employees, and other stakeholders. The IIoT, a vital aspect of Industry 4.0, applies IoT technologies to industrial and manufacturing processes [278]. By enabling real-time data collection and analysis from a plethora of devices and equipment, IIoT enhances efficiency, reduces waste, and optimizes manufacturing processes. Therefore, IIoT plays a key role in achieving the goals of Industry 4.0 by boosting productivity, improving safety, cutting down operational expenses, and fostering innovation.

Additionally, the transition to Industry 4.0 is facilitated by business consortiums comprising technology providers, manufacturers, academic institutions, and sometimes governmental entities [279]. These consortiums collaboratively establish industry standards, share knowledge, and tackle challenges related to cybersecurity threats, data ownership, and workforce evolution. They foster cooperation among organizations to address common problems and expedite the digital transformation journey.

A.1.1 Industrial Data Monitoring

Monitoring is a broad notion that can go from the classic concept of monitoring physical machines up to the more modern and software-oriented concept of monitoring [280]. Furthermore, monitoring has also evolved from an on-site approach to a remote approach due to the evolution of wireless technologies. Even though these concepts are not related at first sight, with the rise of Industry 4.0, there will be a growing number of industrial plants that are based on software and overall informatics-related technologies. Thus, both industrial monitoring as well as IT monitoring will have to coexist. Therefore, in this subsection, we give a few insights on industrial remote monitoring and software monitoring, since, in this work, we cover the monitoring of an industrial environment that includes disruptive IT solutions.

Industrial remote monitoring consists of tracking in real time the data, performance, and security performance of a machine without the user being physically present at the equipment's site [281]. Remote monitoring helps industrial personnel perform a centralized tracking of many machines and even plants at the same time. Specifically, it enables technical personnel to visualize the manufacturing process in real-time by reading data from all the sensors throughout the facility at once. The retrieved information can be combined to have a detailed manufacturing insight. For example, in a filling machine, remote monitoring can track the remaining containers, the machine's actual speed, and how much liquid is remaining. A smart alarm scheme can also be assembled for problem reporting. Finally, remote monitoring can also be used to perform preventive and predictive maintenance. For example, monitoring systems can provide meaningful data regarding lifespan, output efficiency, and breakdown status.

IT monitoring is a complex activity, as many characteristics of many devices must be carefully analyzed to avoid performance degradation. IT monitoring is composed of three sections [282]: foundation, software, and interpretation. The foundation is the lowest part and includes the actual devices and their hardware. The software part includes the monitoring section and includes the analysis of the foundation devices. Finally, the interpretation section includes the gathered metrics, which are presented through graphs, often via a graphical interface dashboard. IT monitoring can be based on agents or be agentless. Agents are independent programs that must be installed on the monitored devices to collect data. Agentless monitoring relies on existing communication protocols to emulate agents, offering similar characteristics as the agent-based approach.

Typically, there are some critical aspects that must be monitored in IT [283]:

- CPU utilization and hardware health and availability.

- Bandwidth consumption between individual devices.
- Firewall and other cybersecurity-related programs, rules, and policies.
- Updates and overall configurations.
- Adherence to basic compliance measures.
- Scalability and throughput.

A.2 Distributed Ledgers

A DLT can be defined as a set of geographically distributed nodes that store and exchange data through a consensus mechanism. In contrast to a classic centralized database, DLTs do not depend on a centralized node, and consequently, they do not have a single point of failure [284]. DLTs generally use P2P technology to exchange data. There are many types of DLTs. Blockchain is currently the most popular DLT since it is the technology behind cryptocurrencies like Bitcoin. In a blockchain, data is organized in "blocks" that are cryptographically linked to each other. However, blockchains tend to be slow and inefficient since consensus algorithms such as the widely used PoW have limited throughput, and high resource consumption [60]. The PoW algorithm effectively avoids malicious behavior in blockchains by requiring the transaction verifiers ("miners") to perform a certain amount of computational effort ("work") in exchange for a reward cryptocurrency. Apart from the heavy computational requirements, another issue is that in a blockchain, every node needs to store a copy of the entire chain, thus requiring each node to possess a significant amount of storage space.

Consequently, novel blockchains and different types of DLTs that intend to replace blockchains have been released. The most relevant and promising solution are DAG DLTs. DAGs were first introduced in [228] with the release of IOTA. In a DAG DLT, the nodes that issue a new transaction must approve two previous transactions and perform a small amount of computational processing to avoid spam in the network. Transactions can therefore be issued without fees, facilitating micro-transactions. DAG DLTs offer huge scalability and throughput, as the more transactions are issued, the faster and more secure the network becomes. Furthermore, the lack of mining makes DAGs highly efficient and suitable for lightweight devices. Thus, this type of DLT is much more suitable for resource-constrained environments that handle a huge number of transactions.

Due to the massive increase in distinct blockchain and DLT platforms over the last years, the interoperability issues have increasingly attracted the attention of the industry [285]. Naturally, there are many different use cases for which different blockchains

have been designed. However, in such an interconnected world, isolated networks are not an option. The use of different blockchains and DLT could be enormously beneficial to take advantage of the latest state-of-the-art technological innovations.

Nonetheless, blockchain and DLT interoperability are not straightforward [286]. In response to this problem, some innovative solutions have been proposed. The most pioneer interoperability-oriented platform nowadays is Polkadot. Polkadot is a highly interoperable solution that consists of a main chain named "relay chain" that governs the network, along with multiple parallel chains that are fully compatible with each other, known as "parachains".

A.2.1 Distributed Ledgers in Industry 4.0

Since the Industry 4.0 revolution started, enterprises have focused on digitizing their manufacturing and business processes. This approach increases efficiency, productivity and profits [24]. However, there are many challenges that need to be solved. These challenges are mostly related to the massive information exchange between a significant number of devices that are geographically distributed. Specifically, some of the most relevant challenges are security, privacy, traceability, and interoperability. Therefore, DLTs have been raised by many researchers and professionals as a possible solution to the aforementioned challenges. The use of DLTs could help eradicate possible single points of failure in industrial networks, along with guaranteeing the integrity of the data and providing traceability of the data from when it is generated up until it is processed at higher levels. Furthermore, smart contracts can provide secure and automated business agreements between various third parties, as well as maintenance and monitoring of industrial machines and processes.

Many big enterprises, such as Amazon, IBM, SAP, Jaguar, DNV-GL, etc., are already exploring the use of blockchain and even DAGs in their business processes. F. Chiacchio et al. [287] demonstrate the viability of blockchain and smart contracts in Industry 4.0 by studying the case of a blockchain-based technological solution for improving the packaging lines of an Italian factory. In this way, the actors that participate in the cycle can retrieve all sorts of information and guarantee the quality of the product.

A.3 Blockchain

A blockchain is a type of DLT [288], in which all the transactions are stored in a chain of blocks that are linked via cryptography, as shown in Figure A.1. The chain continuously grows when new blocks are appended to it. Blockchain provides a distributed software architecture that allows agents (i.e., humans and systems) to interact with each other

without a central authority. In the absence of a central authority, a blockchain network works collaboratively. Each node of the network executes a consensus protocol that defines a set of rules and verification mechanisms to ensure the security, reliability and veracity of the transactions and maximise resilience to failures and cyber-attacks. Specifically, blockchain allows the resolution of conflicts and eliminates information asymmetries by providing a transparent and verifiable record of all transactions, which cannot be altered.

Blockchains can be categorized by the access permissions to their data and by the permissions to participate in the network as a node [62]. Therefore, a blockchain can be permissioned or permissionless and public or private.

- **Permissionless.** In this type of blockchain, all devices can access the network and participate without any permission.
- **Permissioned.** In this type of blockchain, the participation must be authorised, and the actions that can be performed are controlled.
- **Public.** In this type of blockchain, the ledger's information is public for everyone to visualize it.
- **Private.** In this type of blockchain, the ledger's information is private, and only a set of nodes can visualize it.

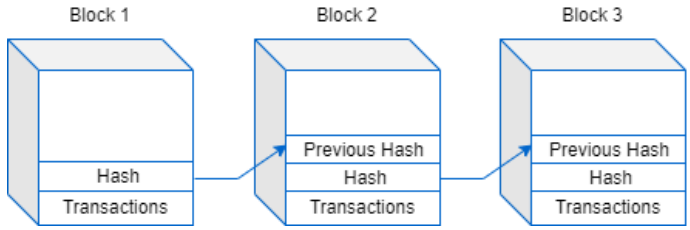

Fig. A.1 Blockchain representation

Each user that performs transactions on a blockchain possesses a pair of public and private keys. The public key is used to provide a unique blockchain address for identification. The private key is used to sign the transactions. When adding a new transaction to the blockchain, the following procedure is followed:

1. The user signs the transaction with their private key.
2. The user broadcasts the transaction to the other nodes of the network.

3. Each peer that receives the signed transaction carries out its validation. If the validation is successful, the transaction is added to each local block that is under construction.

4. When the new block has been completed by reaching the maximum number of transactions that are allowed, or the maximum time established by the blockchain protocol for a new block proposal, the peers acting as miners (i.e., "validators") execute the established consensus protocol.

5. When a miner finishes executing the consensus algorithm, they add the new block at the end of their local blockchain copy.

6. The miner then broadcasts the new block to the network so that the rest of the nodes can verify it. If the validation is successful, then all the nodes of the network add the new block to their own copy of the blockchain so that it remains permanently registered. On the other hand, if the validation is not successful, then the block is discarded.

The consensus algorithm is a key element in blockchain [289]. It establishes the conditions that must be met to reach an agreement between the participating nodes on the validity of new blocks. Ideally, the consensus algorithm should give validators the same vote weight and then make decisions according to the majority of the votes. This scheme may be possible in permissioned networks. However, in public blockchains, this mechanism would lead to Sybil attacks, where a single user with multiple identities (i.e., controlling several nodes) would be able to take over and control the network. In decentralised networks, a user must be selected to add each block. This selection should be done randomly in order to avoid Sybil attacks. The solution proposed by the original PoW-based blockchain (i.e., Bitcoin) [212] avoids such attacks, as it requires miners to perform computationally expensive tasks in order to be elected as validators. Thus, a malicious node would be required to have more amount of computational power than all of the honest nodes. The “work” that is required in PoW-based consensus consists of performing heavy mathematical operations (i.e., mining). Specifically, this process consists of finding a random number (i.e., the nonce) that should cause the hash of the block header to have a certain number of zeros at the beginning. The required number of zeroes is established by a parameter called "difficulty" which establishes how many zeroes are required to be found. The more zeroes, the harder it is to find the right nonce. Despite being very computationally intensive, verifying the results of the mining process is a simple task for the rest of the nodes. However, even though this consensus approach provides great security benefits, it makes blockchain inefficient in

terms of performance, scalability, and energy consumption [290]. Due to the issues mentioned above, several alternative consensus algorithms have been proposed.

The most relevant alternatives to the PoW algorithm are as follows:

- **PoS:** Is a consensus algorithm that requires much less computational power compared to PoW, thus it consumes less energy. In a PoS blockchain, it is assumed that the entities with the highest participation in the network are the least interested in attacking it [291]. Therefore, miners must periodically prove that they have a certain amount of participation in the network (e.g., in the form of cryptocurrency). Because of the advantages of PoS, some popular blockchain platforms, such as Ethereum, are planning to adopt it.

- **PBFT:** Is a consensus algorithm that tackles the "Byzantine generals problem" in asynchronous environments [292]. PBFT assumes that less than one-third of the nodes are malicious. For each block that is added to the largest chain, a leader is selected to validate it. This selection must be supported by at least 2/3 of all network nodes. PBFT can only be used in permissioned networks, and it has a high transaction speed. However, this algorithm has very low scalability, as each network member must constantly communicate with the other nodes. This leads to an increase in the cost of communication as the network grows, making PBFT effective only in networks that have a low number of nodes [293].

- **PoC:** Is a consensus algorithm that allows validator nodes to use their available hard drive space to validate transactions [294]. This contrasts with using computational power (PoW) or stake (PoS). Thus, this algorithm dramatically improves the mining efficiency. However, it has not been implemented in many blockchain networks, and it might enable malware to interfere with the mining process.

- **PoA:** The PoA algorithm consists of a set of trusted validator nodes that approve transactions. Therefore, the validator nodes must not be compromised [295]. In PoA, individuals earn the right to become validators; thus, there is an incentive to retain the gained position.

- **PoET:** Is a consensus algorithm that offers low resource usage and low energy consumption by following a fair lottery system. The algorithm makes use of a random elapsed time to decide the right to mine and the block winners. By running within a secure environment, PoET also boosts transparency by making the lottery results verifiable by external parties [296]. The greatest disadvantage that PoET has is the necessary reliance on specialised secure hardware. Currently, PoET can only be run on Intel CPUs, so the reliance on the consensus model

extends to Intel, a third party. The notion of such a reliance runs against the new paradigm that blockchain networks try to achieve: the complete removal of trust in intermediaries.

- **Proof of Activity (PoAc):** Is a hybrid of PoW and PoS and attempts to bring the best of both algorithms [297]. In PoAc, the validation process starts as a standard PoW process with various miners competing against each other using their computational resources. When a new block is mined, the system switches to PoS. Then, a new random group of nodes from the network is selected to validate the new block. The more stake a validator has, the more chances they have of being elected. In PoAc, a great amount of power is still needed to mine blocks during the PoW phase, and stake owners still have more chances of being elected as validators and accumulating more and more stake rewards. Thus, PoAc combines the best and also the worst of PoW and PoS.

- **RAFT:** Is an election-based consensus algorithm for permissioned blockchain [298]. A leader is elected within the network. The tasks of the leader include accepting node requests and managing the replication of the data. The data flow follows one direction: from the leader to the rest of the nodes. The time in which a leader is in charge is arbitrary.

- **Proof of Burn (PoB):** Is a consensus protocol that requires validators to demonstrate their commitment by "burning" part of their assets in cryptocurrency. The burn process consists of sending funds to an address that can only receive cryptocurrency, but not spend it. The main idea of PoB is that it is better to waste virtual assets like cryptocurrencies instead of wasting real resources (e.g., energy in the case of PoW) [299].

Finally, another key blockchain feature that is worth mentioning is the ability to create smart contracts. Smart contracts were introduced to blockchain by the Ethereum platform. However, the concept of smart contract was first defined in 1996 by Nick Szabo [300] as *"a computerised transaction protocol that executes the terms of a contract"*. Smart contracts are decentralised scripts with sufficient autonomy to be self-executed when certain conditions are met. Smart contracts are included in the blockchain and allow the execution of distributed and highly automated work.

A.3.1 Permissioned - Consortium Blockchains

The firstly released blockchains (e.g., Bitcoin, Ethereum) were public and permissionless. However, these blockchains are typically slow when processing data and are not

scalable. Moreover, they are not suitable for all use cases, such as business consortia. As a response to the aforementioned issues, private and permissioned blockchains have emerged. Enterprise-oriented permissioned blockchains are commonly called "consortium blockchains".

A consortium blockchain represents a form of blockchain network commonly employed when a group of organizations, including businesses or government agencies, need to exchange information and collaborate on shared objectives while retaining a certain control over network access and validating process [277]. This type of blockchain contrasts with public blockchains such as Bitcoin or Ethereum, where the ledger is accessible to everyone, and anyone can become a validator.

Consortium blockchains typically employ the same foundational technology as public blockchains, such as cryptography for securing the network and a distributed consensus mechanism for transaction validation. However, consortium blockchains usually employ a different consensus mechanism compared to public blockchains. Due to their permissioned nature, consortium blockchains typically use lightweight consensus mechanisms such as the PBFT [292] consensus, which is designed to be more efficient and faster than the PoW algorithm that is commonly used in public blockchains such as Bitcoin. PBFT works by having a designated leader, known as the primary, that is responsible for ordering and broadcasting transactions to the rest of the network. The other nodes then verify the transactions and reach consensus on their validity. PBFT, however, has relatively low scalability since there is a great amount of constant communication between the nodes. Thus, the more nodes are in the network, the slower it is. As a result, alternative permissioned consensus algorithms such as Raft have emerged [301]. Raft uses a leader-based approach, where all writes go through the leader, and all followers agree and apply the writings. This allows for a simpler and more understandable algorithm while still providing safety and liveness properties.

A.3.2 Permissionless - Public Blockchains

As mentioned before, early blockchain platforms such as Bitcoin and Ethereum introduced the concept of permissionless, public blockchains. However, their slower transaction processing speed and scalability issues made them less suitable for certain applications, leading to the development of scalable public blockchains.

These scalable public blockchains maintain the open-access principle of traditional blockchains, allowing anyone to access the ledger and participate in the validation process, while also addressing scalability limitations. They employ advanced consensus mechanisms like PoS and Sharding, instead of the slower, energy-intensive PoW algorithm.

PoS chooses validators to create a new block based on their economic stake in the network, allowing for faster and more energy-efficient transactions. Sharding, meanwhile, divides the blockchain network into smaller parts, each capable of processing transactions and smart contracts independently. This parallel processing increases network capacity and speed. Ethereum's 2.0 upgrade is one implementation of this sharding mechanism.

In summary, scalable public blockchains offer the transparency and openness of original public blockchains, coupled with improved scalability, making them a significant evolution in blockchain technology. However, public-permissionless blockchains are still a poorer alternative when compared to private-permissioned blockchains in many use cases.

A.3.3 Blockchain and IoT

The IoT can be defined as the interconnection of everyday objects that are connected to the internet. One of the core features of IoT is linking the physical world and digital world together. Sensors play a very important role in IoT [302]. Sensors collect data from the environment, which generates a great amount of useful information. According to [303], the development of IoT includes three phases: embedded intelligence, connectivity and interaction. Embedded intelligence means that devices can perform actions automatically. Connectivity in IoT is mostly given by wireless connections such as ZigBee, WiFi, 3G, etc. Finally, IoT devices must also be capable of interacting with each other autonomously. Thus, with IoT, the current human-to-human interaction will turn into machine-to-machine interaction. The identification of IoT devices is made mostly by the use of RFID. RFID is an extension of the optical tags that are found in everyday objects. These tags include embedded intelligence so the identity of an object can be decoded remotely [304].

The IoT generates large volumes of data and requires connectivity and power for long periods of time [305]. This, together with the limitations of the network, computational capacity and limited power supply lead to a high number of challenges. Furthermore, heterogeneity in IoT networks is currently too high due to the lack of standard protocols in this field [306]. Other crucial challenges of IoT are privacy and security. In the current centralised IoT architectures, we cannot be sure if the data has not been tampered with, altered or falsified. Also, nowadays, in many areas, the traceability of assets during their life cycle is required, thus making the immutability of the data a key challenge.

Blockchain is considered by many researchers as the most appropriate solution to the challenges that are present in IoT due to its key features such as security, immutabil-

ity, trust and decentralisation [307]. Blockchain could protect IoT networks against data tampering. Furthermore, the possibility of creating automatised software that is shared over a decentralised and cryptographically secure blockchain network would increase the autonomy of IoT. In addition, the lack of a central authority would make IoT able to operate more quickly. Furthermore, decentralisation would eliminate single-point failures, thus improving the security and reliability of IoT. The immutability of blockchain is also ideal for the traceability of the data.

A.3.4 Smart Contracts and Oracles

Initially, blockchains could only process simple transactions. Consequently, in 2015, the Ethereum project introduced the execution of smart contracts. A smart contract is a program that is executed on top of a blockchain network. With smart contracts, the blockchain has greatly expanded its range of applications from simple financial transactions to more broad and complex applications in industry, smart homes, healthcare, etc. In the field of industry, blockchain and smart contracts could be used to establish automated and trustworthy agreements between different business partners, clients, and suppliers and increase the confidentiality, privacy, and security of IIoT data [24]. Nonetheless, many smart contracts require external information to make decisions properly. Besides the fact that off-chain data could be challenging to be accessed by a smart contract code, external dependability could also undermine the advantages of blockchain networks by removing decentralization and trust [308].

To solve the aforementioned issues, the concept of oracle has been introduced. In computer science, an oracle can be defined as a service that provides reliable data from outside a specific system [309]. However, centralized oracles introduce a single point of failure within blockchain networks. This issue might lead to the introduction of corrupted data inside smart contracts, which would compromise the whole blockchain network and make it pointless in terms of the security of the information. Thus, blockchain oracles are needed. A blockchain oracle is a decentralized oracle that is capable of analyzing the external world and providing trustworthy data to smart contracts [310]. Currently, there are many blockchain oracle services in the market, the best-known being ChainLink. ChainLink enables simple deployment of decentralized oracle networks that are capable of interacting with the Ethereum blockchain via Solidity smart contracts. Apart from ChainLink, there are other relevant solutions such as Augur, which is mostly focused on decentralized finance, and Gravity, which claims to be highly efficient and secure, but is in a too early stage of development [311].

However, taking into account the definition of oracle, we can state that any blockchain can be used as an oracle service. Nonetheless, oracle-oriented blockchains offer greater

smart contracts compatibility and their software implementation graphical interface is oriented towards oracle monitoring. For example, most blockchain oracles incur economic costs, thus needing strict monitoring to maximize their efficiency and reduce costs. However, oracle-oriented services may have limited compatibility with other services or limited functionalities [312].

A.3.5 Blockchain Interoperability

Blockchain interoperability refers to the ability of distinct blockchain networks to communicate and interact with each other. This can be achieved through intermediary gateways or through the use of compatible parallel chains [313].

In parallel chain architectures, a leading chain governs the entire network and offers out-of-the-box compatibility between the rest of the parallel chains that compose the network. Each parallel chain serve a distinct purpose, and includes it unique features and capabilities. Moreover, these types of protocols typically support different types of consensus mechanisms and can be used to create private or public networks. Even though parallel chain protocols offer a straightforward, simple approach to interoperability, they are still limited when interacting with other blockchains that are built by different companies. Furthermore, parallel chain approaches are typically complex. Currently, the most known parallel chain platform is Polkadot [314].

Gateway-based interoperability is a method of connecting different blockchain networks by using an intermediary gateway connector that acts as a bridge between them. This approach allows for the transfer of assets and data between the different networks, enabling interoperability [286]. A gateway can be implemented in many ways, from a simple API connector to more complex systems such as intermediary node networks, to blockchain oracles that are programmed for interoperability tasks. A primary benefit of gateway-based interoperability is its ability to facilitate asset and data transfers between distinct blockchain networks without necessitating intricate mechanisms or extensive cooperation between various blockchain providers. Nonetheless, gateway-based interoperability has its drawbacks, such as the presence of centralized points of failure. Consequently, if the gateway network is compromised, all dependent networks may be impacted. Despite these drawbacks, gateway interoperability remains a straightforward and efficient approach to achieving interoperability.