Towards Decentralized Operationalization of Zero Trust Architecture

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

05-02-2022 / 10-02-2022

CITATION

Pokhrel, Shiva Raj; Li, Gang; Doss, Robin; Nepal, Surya (2022): Towards Decentralized Operationalization of Zero Trust Architecture. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.19127090.v1>

DOI

[10.36227/techrxiv.19127090.v1](https://doi.org/10.36227/techrxiv.19127090.v1)

Towards Decentralized Operationalization of Zero Trust Architecture

Shiva Raj Pokhrel, Gang Li, Robin Doss and Surya Nepal

Abstract—Zero Trust Architecture (ZTA) is a new security concept that mainly relies on the five tenets: situational awareness, dynamic authentication and authorization, fine-grained access control, small trust zones and continuous trust evaluation. ZTA is increasingly popular every day, but operationalizing it in emerging networking contexts such as working from home (WFH) has been highly challenging and non-trivial. In this article, we are the first to propose ZTA operationalization approaches into three distinct categories: i) centralized operationalization of ZTA (cZTA), ii) federated operationalization of ZTA (fZTA), and iii) decentralized operationalization of ZTA (dZTA). cZTA can be viewed as a framework where a trusted central server calculates the trust values for every device in the system (e.g., single sign-on, multi-factor authentication etc.) but has several problems. We proposed fZTA, ZTA operationalization by extending Google’s federated learning (FL), which can address the authentication problem of cZTA. Every other aspects of trust and security are still based on the centralized framework. To facilitate secure and trustworthy WFH and collaboration for enterprise networks, we develop a novel operationalization model, dZTA, by leveraging ZTA’s concept and adopting decentralized computing capabilities from the distributed ledger and FL. New insights are the development of dZTA framework by implementing i) distributed policy engine modelling, ii) policy enforcement design, iii) decentralized trust computation and iv) new information leakage prevention strategies, for the decentralization of identities. By considering the core tenets of ZTA, we implement the proof of concept of dZTA over WFH in enterprise networks as a case study to demonstrate its feasibility.

Index Terms—Decentralized Operationalization Approach, Federated Learning, Zero Trust Architecture.

I. INTRODUCTION

Uncertainties are unavoidable in today’s enterprise and remote collaboration networks. COVID-19 has accelerated digital transformation strategies, including the 5G technologies and the IoT, along with the unforeseen growth in remote working and collaboration of devices [1], [2], [3]. Companies and organizations moved rapidly into *working from home* (WFH) to sustain team productivity and implemented automation and orchestration of their networks to minimize business disruptions. Security and trust in such evolving enterprise networks should never be taken for granted to reduce uncertainties and mitigate risks but continually be assessed.

With the modern advances, remote communication and control mechanisms are easy to maintain, which provides users/devices with easy access anywhere, anytime, and on the move. However, one primary challenge to realize the full potential of the remote working, automation and orchestration

paradigms, is the *security architecture* in critical infrastructure [4]. Cybersecurity architecture is certainly not immune to this significant change. Existing security approaches cannot deliver the required (or desired) levels of security protections.

Managing WFH access, remote workforce monitoring, and traffic analysis are among the foremost problems. The increasing trend towards high mobility, remote working, WFH, and distributed collaboration of IoT applications has challenged the previous notions of network security, especially the perimeter-based security [5]. The concept of zero trusts [6] came around, because the seminal perimeter-based network security model of ‘*inside means trusted*’ and ‘*outside means untrusted*’ no longer works. This perimeter-based approach has been tweaked over the years using *virtual private networks* (VPNs) and *demilitarized zones* (DMZs) to deal with new challenges, such as users/devices becoming mobile and business partners (and devices) outside the network that need access. However, an inherent drawback exists in today’s highly complex enterprise networks – *they grant excessive implicit trust*. Once connected, whether directly or using a VPN, we are then trusted everywhere inside the internal network and applications.

A. Related Works

In recent years, real momentum towards a highly distributed workforce has exposed the weaknesses of established perimeter-based security approaches (e.g., enterprise firewalls) [7]. It has already been noticed that perimeter-based access fail to: 1) effectively mitigate malicious actors/devices inside the network perimeter, and 2) offer a sufficient level of protection for users/devices and assets located outside the defined network perimeter (e.g., remote users, devices, services, applications.) Therefore, robust security models, entirely independent of perimeter and location are much-needed for future networks [7].

The *Zero Trust Architecture* (ZTA) [8] is a significant departure from perimeter-based security models. A brief history of ZTA can be observed from the works [8], [9], [10], [11], [12], [13] and the references therein. After the origin of zero trust [8] and key initiatives from *Google* [9], in recent years, momentum towards standardization of ZTA concept (e.g., *Gartner* [10], [12] and *NIST* [13]) has been building up gradually [14].

As shown and explained in Fig. 1, ZTA proposes a radically different approach to network security, in which the infrastructure is always assumed compromised¹. The ZTA paradigm shifts security away from implicit trust dependent on a network location.

S. R. Pokhrel, G. Li and R. Doss are with the Centre for Cyber Security Research and Innovation (CSRI) at Deakin University, Australia and S. Nepal is associated with Cybersecurity Cooperative Research Centre (CRC) and CSIRO DATA61, Australia.

¹Fig. 1 is motivated with relevant insights from the ideas proposed by Microsoft tools and services offered over Azure.

Figure 1: ZTA advances security well beyond the typical perimeter-based strategy (AKA company firewalls). The basic idea of ZTA is that organizations should not trust anything within or outside their perimeter(s) automatically and instead check anything or all that attempt to connect to their systems before giving access to them. Therefore no access to any resource should be given until we know who is interested and given the minimum privilege principle is always satisfied.

ZTA relies on the five tenets: *situational awareness, dynamic authentication and authorization, fine-grained access control, small trust zones* (micro-segmentation) and *continuous trust evaluation*. It follows a basic principle – ‘never trust, always verify,’ and applies it to both users and assets at an utmost level of fine granularity. The three main reasons why corporates should consider adopting ZTA for networks security are: *trust through device verification, improved data protection and continuous authentication-cum-authorization*. Overall, ZTA presents a promising vision for the required security that is focused on users, assets, and resources – rather than purely on network access.

Rather than a network location or IP address, ZTA mandates that the trust need to be derived directly from a mixture of identity and context-based elements. In sharp contrast with perimeter-based, the primary goal of ZTA is to minimize the uncertainty associated with the enforcement of accurate and least-privilege access decisions per request [12], [13].

Nevertheless, operationalising ZTA in today’s enterprise networks is a highly challenging and non-trivial task. In other words, how to realize the above mentioned five tenets effectively and securely in a complex enterprise or cyberphysical infrastructure networks without trading-off resource availability or dynamic control over network assets is an open problem that requires novel solutions. Moreover, due to WFH devices, the risks presented by this increased and diversified attack surface and the vulnerabilities have created the need for operationalising ZTA that is fully *decentralized, intelligent, agile, and highly responsive* [15], [16], [17].

Current approaches to ZTA, including the *Gartner’s* extended approach [12] and the *National Institute of Standards and Technology* (NIST) operationalization framework [13] are fully centralized, presenting both security (e.g., *single point of failure, denial of service*) and performance (e.g., scalability, speed) challenges [18]. In particular, NIST provided a conceptual framework with the following two primary objectives: 1) preventing data breaches, and 2) limiting internal lateral activities. Along the lines of NIST vision, the main aim of this research work is to identify techniques for developing a comprehensive and concrete framework of a *decentralized operationalization of Zero Trust Architecture* (dZTA). The proposed dZTA aims for trusted remote working and collaboration that is scalable and efficient, provably secure, and optimized for preventing

information leakage while limiting the opportunity for lateral activities by both malicious and non-malicious actors.

B. Contributions

Main contributions of this work are the detailed guidelines towards developing the components of the proposed ZTA operationalization methods over enterprise networks. Our main contributions are:

1. We propose to classify ZTA operationalization approaches into three distinct categories: i) centralized operationalization of ZTA (cZTA), ii) federated operationalization of ZTA (fZTA), and iii) decentralized operationalization of ZTA (dZTA). We demonstrate that fZTA (operationalizing the ZTA concept by extending Google’s federated learning (FL)) can address the authentication issues. But, every other aspect of trust and security of WFH networks are still based on the centralized framework. Moving ahead, we develop a new decentralized framework for operationalizing the ZTA concept (i.e., dZTA) by considering the decentralization of identities. Our approach to classify the operationalization methods of the ZTA concept is explained in Sec. II). To this end, we developed a thorough understanding of the ‘network security needs’ for remote working and collaboration as well as the need and design gaps towards a fully decentralized operationalization of ZTA. The ideas proposed in this paper are both novel and original as one of the first attempts towards operationalizing the ZTA concepts.

2. We design a holistic decentralized operationalization framework for the enforcement of ZTA design goals. The novelty here is to enhance the accuracy of the decentralized learning and data analysis of the subjects/devices for policy enforcement, while relaxing the dependency on the trusted central curator. Moreover, by developing and evaluating efficient and incentive-compatible (fully) collaborative learning algorithms for trustworthiness and rewarding, the most audacious aspect of the decentralized operationalization of ZTA lies in the real-time remote collaborative system within the performance, security and data utility trade-off derived from the findings.

3. New outcomes and findings are the development of dZTA framework by implementing i) distributed policy engine modelling, ii) policy enforcement design, iii) decentralized

Figure 2: ZTA operationalization has started evolving from centralized to federated approaches. We employ Federated ZTA (fZTA) as an approach for ZTA operationalization by extending Google's FL [19], and can only addresses authentication issues—every other aspects of trust and security of networks are still based on the centralized framework. Beyond fZTA, we propose a new operationalization advancement for ZTA considering decentralization of identities, called dZTA, which is another main focus of this work.

trust computation and iv) new information leakage prevention strategies, for the decentralization of identities.

Our dZTA focus on accessing resources in a secure decentralized manner regardless of network location, users, and IoT devices. It enforces rigorous access controls by continually inspecting, monitoring, and logging network traffic. We anticipate the need for designing novel data-level protections, robust identity architecture, and strategically distributed learning algorithms to build trust zones in enterprise networks and around corporate resources (recall Figure 1).

II. ZTA OPERATIONALIZATION APPROACHES

We classify ZTA operationalization approaches into three categories: i) centralized operationalization of ZTA (cZTA), ii) federated operationalization of ZTA (fZTA), and iii) decentralized operationalization of ZTA (dZTA). In Figure 2, we have compared and explained them in details. Centralized ZTA operationalization refers to the framework where a trusted server calculates the trust values for every device in the system (e.g., single sign-on, multi-factor authentication etc.). All devices of the system request this server to provide them with information about other devices. Federated operationalization of ZTA, for example, is an improved ZTA operationalization by extending Google's FL [19], which can address the authentication problems of cZTA – other aspects of trust and security of networks are still based on the centralized aggregations of models [20]. These two operationalization schemes, viz, cZTA and fZTA, has severe bottleneck problems:

- each device depends on the trustworthiness of the single server, i.e., a *single-point-failure*;
- different devices may have different opinions about the trust of any device, however, such a fact cannot be considered.

The decentralized operationalization technique, dZTA, proposed in this paper corresponds to each device being the “center of its own world.” In particular, devices are responsible

for evaluating their trust values themselves (using FL and blockchain) for any target they want to associate with. Such a novel idea is motivated by the “bottom-up” approach that has been successfully implemented and adopted, as a part of *Pretty Good Privacy* (PGP) [21] for public key certification. It is worth noting that the difference mentioned above quantifies the semantics of the trust. For dZTA, constant collaboration is often required by subjects, and most of the cybersecurity measures need to be facilitated efficiently in such collaboration. Perhaps the most efficient method is by providing federated identity management systems that allow users of one company to utilize their verifiable identities across agencies. To this end, we anticipate that such decentralized identity management across enterprise boundaries can be achieved by dZTA.

A. Why enterprise network security requires decentralized operationalization of ZTA?

It is worth recalling that ZTA is a concept in which no link is trusted unless explicitly authorized. ZTA attempts to protect every user, every device, and every connection at all times. To safeguard our most precious assets and handle threats more effectively, ZTA can combine and integrate our security technologies. This is reflected, for example, by the recent moves of IBM towards putting ZTA solutions and services into action to fuel business without limits². ZTA can be implemented by segmenting the network following corporate security requirements and then realizing safety regulations for orchestrating the underlying network communication between these segments. Today's enterprise networks are naturally distributed and function in decentralized fashion. Digital transformation and growing IoT and cloud adoptions are creating increasingly dynamic corporate ecosystems. The cybersecurity market is packed with points that usually perform one thing

²<https://www.ibm.com/au-en/security/zero-trust>

and only one thing on a specific OSI layer to address corporate security demands, which leads to several problems, including i) policy across products must be consistent, ii) products communicate and interact without generating policy deviations; iii) their security events must be correlated, among others.

Limitations of cZTA. As mentioned, today's corporate networks operate in a decentralized fashion. Therefore, a central segmentation gateway between devices/users and the data center network is required for operationalizing ZTA. Such a centralized method can be generally utilized for WFH devices/users who use data centre resources as a replacement for VPN and may thus be extended for office and working from home setup. A cZTA system at a gateway can only cater for one vector - malware from a device/user into the data centre network. As a required intermediate gateway, it can neither operationalize ZTA between WFH devices and their interactions in the enterprise environment nor impose ZTA between servers' interactions.

In addition, for internal office users, the centralized portal strategy frequently involves a network change effort to divide the existing flat business network. Therefore, users at the office need to proxy the data center server via the gateway, and the changes to routing, DNS, and firewalls are always mandatory. This often causes a bottleneck in the gateway that requires many clustered gates to reduce performance and faults. In addition, the context for the apps requires precise inspection capabilities (to separate web-based access and database access), which influences performance further. The data centre servers must now be fully accessible through the gateway, thus, breaching the 'never trust - always verify', one of the main ZTA design goals.

dZTA: the way forward. A decentralized peer-to-peer operationalization approach can be used to implement ZTA concept over enterprise networks. This works particularly well on the flat corporate and WFH networks, where resources may connect directly without any gateways. A host-based firewall enforces endpoint policies (servers or IoT devices) that are organized into a central policy orchestrator using local knowledge of these endpoints. The agent, which resides on the system, delivers application-level situational awareness to the network flows, which is in sharp contrast to cZTA (and therefore, dZTA are not subject to precise inspection at the application level). The dZTA is resilient and functions well for campuses and WFH networks, and data centers as well as between office users and the corporate houses.

Pros of dZTA. dZTA facilitates natural zero-trust implementation over remote working and collaborative enterprise networks. The scope and pace of this transformation will be more efficient and fully under the control of the enterprise as it happens at the speed and scale of their network. This is in sharp contrast to cZTA, and the network never requires redesign for ZTA operationalization. For instance, with dZTA, the data plane is never failing at any single point - no single point failures.

B. Situational awareness for ZTA

It is worth observing that leveraging situational awareness for continuous diagnostics and monitoring requires adopting

the advances from machine learning. In addition to this, realizing the proposed dZTA at the necessary scale and speed to support large-scale WFH and collaborative settings demand efficient distributed learning mechanism. To the best of our knowledge, the most suitable distributed learning framework for capturing the situational awareness for continuous diagnostics and monitoring is the FL [19].

However, in current FL, a central server orchestrates the collaborative learning process and receives the contributions of all devices. All of the existing distributed learning frameworks, including Google's FL [19], require global coordination using a centralized server, thus resulting in federated operationalization of ZTA, i.e., fZTA. Such coordination procedures through a single centralized server often incur long delays, and are unreliable (e.g., *single point of failure*, *denial of service*). See other details of fZTA and its limitations in Figure 2. In addition, any compromise of the local learning model or data poisoning can be often cascaded, potentially causing deception of the global learning process. We also find that integration of the recent advances in distributed ledger techniques and decentralized learning frameworks provide new opportunities for the decentralized operationalization of ZTA. Blockchain-based FL is the best alternative for facilitating dZTA involving remote working and WFH collaboration by accounting the decentralization of identities. [15], [22]. For instance, by leveraging the celebrated properties of Blockchain over FL, such as immutability, tamper resistance, and traceability, dZTA becomes feasible with blockchain-based FL [23], [24]. Moreover, FL with Blockchain enables operationalizing the tenets of a ZTA and implementing situational awareness in a decentralized fashion. We anticipate that dZTA by ablating a few FL components with successful Blockchain elements can facilitate the desired dZTA over trusted remote working and WFH collaboration of enterprise networks (c.f. Figure 2).

C. Decentralized Operationalization Ideas

After we implement the dZTA and capture the situational awareness, the learned situational awareness is to be leveraged for trust computation as well as continuous authentication and authorization purposes.

Such a dZTA has the advantage of continually exchanging local learning updates of the situational awareness from all WFH devices and newtrk devices while verifying (and providing) their corresponding accesses and authorities autonomously (using the distributed ledger). The design of a lightweight distributed ledger-based dZTA requires the careful design of a decentralized FL framework to achieve robust and fully decentralized learning - a core component of the envisioned dZTA.

To this end, we have identified the following three fundamental research ideas:

1) *Decentralized trust computation:* Fundamental to dZTA for secure remote working and WFH collaboration is the secure and reliable trust computation mechanism. Trust establishment in a WFH context, where the context of the user can dynamically vary in terms of physical location, device use, network location, behaviour, etc., requires novel algorithms for

trust computation that can dynamically adapt to new contexts through continual incremental learning.

2) *Decentralized learning and orchestration*: In the proposed dZTA, we need to achieve remote collaborative learning for situational awareness without reliance on a central server (and any other centralized orchestration), which is a significant departure from current FL approaches and fZTA [24], [25], [26]. A fully decentralized operationalization approach addresses the issue of a *single point of failure* carried over to the existing FL frameworks, while also providing an opportunity for privacy-preserved data sharing and learning. Another limiting factor to be considered here is the limited computing capabilities and communication constraints of the remote WFH devices.

3) *Prevention of information leakage*: Beyond the core ZTA tenets, information leakage issue on distributed data exchange emerges in the proposed dZTA. This results due to an increasingly dynamic ecosystem of integrated information systems, applications and WFH devices in enterprise networks. The total incorporation of privacy controls into protection has been one of the significant objectives of the next generation NIST SP 800-53 (rev 5) privacy and security architecture (Joint Task Force Inter agency Working Group 2020) [13]. Tracking information leakage in a large-scale network is a daunting task, and the accurate tracking over WFH system dynamics and accountability have non-negligible limitations in practice – high overhead and complexity [27]. With the proposed dZTA for remote and collaborative WFH, the development of fundamentally distributed approach to mitigate dataset level attribute leakage and will become more tractable in the blockchain-based FL setting [28].

III. PROPOSED DECENTRALIZED OPERATIONALIZATION OF ZTA

Centralized operationalization of ZTA is quite straightforward based on the NIST standard [13] guiding security principles directly for the workflows, security design, and operations. A high level conceptual framework for cZTA employing NIST architecture is shown in the left panel of Figure 3, which illustrates the fundamental relationship between different ZTA components and their interactions [13]. In this ideal ZTA model, with different logical parts and their interactions, we can see that the policy decision point of ZTA is decomposed into three modules: *policy engine*, *administrator* and *policy enforcement point*. The first two (*policy engine* and *administrator*) operates at the control plane to communicate, while policy enforcement happens on a data plane. With relevant insights drawn from NIST ZTA concept, based on a hierarchical architecture [29], our design dZTA consists of two main parts as shown in the right panel of fig. 3.

We develop the decentralized policy engine in our framework, and the modelling of decentralized policy enforcement are based on the five tenets of ZTA concept aimed at a distributed computing WFH environment. As shown in fig. 3, we ablate FL with Blockchain for autonomously tracking situational awareness and guaranteeing the dynamic authentication and authorization at the control plane. With fine-grained access

control for small trust zones, we develop a mechanism as shown in fig. 3 for the continuous trust evaluation by integrating trustworthiness with reward mechanism at the data plane. The details of fig. 3 are explained as follows:

a) *Developing the decentralized policy engine*: *Decentralized Policy Engine* (DPE) is essential for the final decision of granting/revoking access to a resource for a given case and can be executed at the local devices [30]. With a fully distributed ledger [31] and FL [32] capabilities, DPE can use enterprise policy to evaluate the decision to grant or revoke access to the resource [24]. When coupled with the administrator component, DPE generates and logs the decision while the policy administrator executes it.

b) *Modelling decentralized policy enforcement point* (DPEP) with trustworthiness and reward mechanism: Connections between a WFH user device and an enterprise network resources are enabled, monitored, and finally terminated by DPEP. In contrast to cZTA, our design split this logical component into two parts: i) the FL agent on the WFH user's device, and ii) the blockchain-based gateway component in front of resource that controls access. DPEP operates in the data plane and as shown in Figure 3 the proposed DPEP consist of three primary implementation levels, viz. 1) Data/information leakage control layer; 2) Blockchain-enabled decentralized learning; and 3) ZTA policy enforcement layer.

IV. dZTA DESIGN DETAILS

DPEP in the proposed dZTA works as follows. Collaboration or compilation of all observed data from either the local area or other manual entry details via different sources requires the prevention of data/information leakage. The transactions that record the collaboration and communication would be stored in blocks; However, the observational data for ZTA are to be kept out of the chain for computing a hash function. Blockchain-based FL operates in between the data and policy enforcement layer to enable two-way (bidirectional) communications. Furthermore, with dZTA, as the whole enterprise network can not considered an implicit trust zone, blockchain-based FL facilitates to create *small trusted zones*. Resources must continuously operate considering that there is an attacker on the enterprise network, and communication must be performed in the most secure way possible. FL can guarantee this by enabling small trusted zones because data never leaves the device or resources. Furthermore, in dZTA, we execute authentication of all connections in each FL communication round and embrace local and global information exchange (encrypting raw data traffic using machine learning models).

As shown in Figure 3, we develop two necessary modules to establish the desired level of trust and maintain a rewarding mechanism. Furthermore, we explain the three core components for dZTA and how they interact with each other to attain the *five tenets of ZTA concept* in the following subsections.

A. Continuous Authentication/Authorization with Collaborative Intelligence

In order to be truly successful, ZTA requires a component to continuously monitor the remote WFH applications, and

Figure 3: Proposed decentralized operationalization of ZTA (dZTA) framework and interrelationships among different components.

collaborative APIs. In the considered network setup, WFH users are accessing and pairing that visibility with their identity solution using FL based local models. This tight coupling of FL mechanism with authentication and authorization performs the desired continuous authentication and authorization in each of communication rounds for federated aggregation. Furthermore, when any anomalous behaviour occurs, dZTA framework supports triggering re-authentication and authorization (as and when necessary).

Authentication and permission for one resource does not allow access to another resource automatically. Besides dZTA performs periodic re-authentications via native communication rounds during federated aggregation, the key is to know when a re-authentication is necessary due to some malicious or simply anomalous event taking place. We have a dynamic policy enabled by the blockchained-based FL, which continuously acts as an interface for learning the observable state of the client identity, application, and the requesting asset, as well as other behavioural factors for actuation which determines access to resources. This is where the combination of identification/authorization technology with remote application and API protection comes into play. Therefore, with relevant insights from this work, we need to develop a novel context-aware

continuous remote authentication protocol empowered with intelligent collaborative learning.

B. Decentralized Policy Engine and Administration

An orthogonal question for the proposed dZTA is how can it be practically realized? With relevant insights from the ZTA [13], our idea is to ablate the successful feature of distributed ledger with FL, which is explained next.

FL has been known to leverage decentralized personalized data from a set of WFH devices/users and holds the potential to be an enabler for the anticipated dZTA design. FL can significantly improve the execution while ensuring local computation for the ZTA. It has the potential to exploit widely distributed training of data, while maintaining local learning within each device or a network edge.

Yet, such a mechanism incurs long processing delays and risks due to a single central server. The central server aggregates local updates for a global model by assembling local updates from endpoints in every iteration. However, any unreliable model uploaded by the user/device (if any) often leads to deception in the FL. Some fraud devices may perform unreliable updates intentionally, e.g., the data poisoning attack, or unintentionally, e.g., low-quality data caused by energy

constraints or high-speed mobility. Therefore, investigation towards a trustworthy and fully decentralized learning approach for the remote and collaborative working becomes critical.

To this end, we divide this component into two: 1) Lightweight blockchain-based decentralized learning framework, and 2) Decentralized Policy Enforcement.

1) *Lightweight Blockchain-based Decentralized Learning Framework*: In this subtask, at the heart of the dZTA operationalization design we integrate trustworthiness and reward modules (cf. Figure 3), which learns and jointly evaluates the trustworthiness of end-points observations during model aggregation and adapts the block verification promptly at the Blockchain level. Such a new framework for remote and collaborative working is to jointly evaluate device observations' and trustworthiness with FL aggregation.

We develop an approach to adapt block verification at the Blockchain level via the coupling of trustworthiness and reputation algorithms. Such a new mechanism is to deploy dZTA efficiently over the remote WFH environment and to provide a new scheme for secure collaborative working solutions. Observe that one of our key objectives in this design is to prevent unwanted access to data and services while also ensuring *granular access control* enforcement. To put it another way, only authorized and approved subjects (a mix of users, applications, and devices) have access to the *granular blocks* in the proposed dZTA. Other subjects such as attackers can not access the blocks. Looking ahead, as indicated by NIST [13], we are considering resource access as a whole (e.g., printers, computing resources, WFH devices, data, models, etc) in the form of a granular block of dZTA rather than only the data access.

The DPE can be regarded as the brain of a ZTA operationalization and the PE's trust algorithm as its fundamental cognitive process. To grant, deny, or revoke access to a resource, the DPE employs enterprise policy as well as information from external sources (e.g., CDM systems, threat intelligence, information leakage prevention services discussed below) as inputs to the trust algorithm. The trust algorithm determines whether or not to give or deny access to a resource. The DPE gets information from various places: i) The policy database, which contains observable details on subjects. ii) Subject's traits and roles. iii) Past experience, behaviour patterns. iv) Threat intelligence sources. v) Information leakage prevention and other metadata.

2) *Information Leakage Prevention* : Continuous trust evaluation in remote WFH and collaborative working involves data sharing and information flow across different locations, and the collaborative learning (or analysis) relies on the information (data and domain knowledge) accessibility from the data miners. This full or partial transfer of ownership of information, if not properly carried out, may lead to leakage of information, which refers to the disclosure of essential information for which data owners are not willing to disclose, including, but not limited to, private data of users, enterprise policy, business secret information, etc.

Our first main step for formal treatment is the thorough characterization of various actors and their duties, which will eventually determine suitable privacy designs. For example, it is important to differentiate the server administrator's view

from the view of data miners that induce learned models. It is possible that a tool designed to continually provide strong privacy protections against a malicious data miner may not provide sufficient security for an adversarial system or malicious server.

Typically, we need the composition of several tools and technology into a trustworthy remote end-to-end working system, all layering several strategies to safeguard the system as a whole (e.g., developing procedures for executing portions of a *Secure Multi-Party Computation* (MPC) protocol inside a *Trusted Execution Environment* (TEE)). This should be making it intractable for the adversary to compromise the component sufficiently and redefine different strains of information to avoid leakage for the purpose of FL and other components in the dZTA framework, e.g., developing a new MPC to protect the aggregation of model updates, then using private data releasing/disclosure techniques before sharing the aggregate updates beyond the server [33], [34].

As such, in this component we need to build new federated algorithms for remote working systems wherein the privacy properties degrade as gracefully as possible, in cases where one technique or another fails to provide its intended privacy contribution. For example, by executing the server component of an MPC protocol inside a TEE allows privacy to be maintained even in the case where either (but not both) the TEE security's or MPC security assumptions fail to hold in practice. This requires further investigations in future.

3) *Decentralized Policy Enforcement Point (DPEP)*: In this subtask, the policy enforcement will receive the information from decentralized learning for ultimate authentication/authorization to the devices/users. Depending on the user/device requirements, the dZTA policy enforcement could potentially adapt the block validation mechanism. The confidence of the policy enforcement is anticipated to be quantified by the trust module based on 1) the evidence, and 2) the long-term behaviour, the reward of the nearby devices/users or sources.

The long-term trust of network participants can be tracked by the reward module (using blockchain). It will take inputs (information based on the participant's reward history) from the decentralized learning level and continuously (re)compute the reward value. Such updated reputation values can be used dynamically by both decentralized learning and policy enforcement.

V. PROOF-OF-CONCEPT: dZTA FOR WFH NETWORK

In this section, we develop proof-of-concept of dZTA for WFH settings using FL with Blockchain and evaluate the developed ideas.

A. Network Settings

Given N WFH devices and M data miners, we consider $n = 1, 2, \dots, N$ remote working devices in a network system, where the n^{th} device fetches a set of trust samples t_n , and performs its local trust learning update. The updated learning of local workers from the n^{th} device is then transmitted to a data miner m . This association from the devices to data

miners is considered to be randomly uniformly chosen from $m \in \{1, 2, \dots, M\}$, as shown in Figure 3.

In this context, our FL in dZTA is a regression, which solves the problem by accounting the whole confidence state space of trust samples, $T = \cup_{n=1}^N t_n$ fetched by all devices in our dZTA network system. Considering a global trust vector G_T , the primary goal is to

$$\min \{F(G_T)\}, \quad (1)$$

$$\text{for } F(G_T) = \frac{1}{|T|} \sum_{n=1}^N \sum_{z_j} (x_j^\top G_T - y_j)^2 / 2$$

for convenience, $z_j = \{x_j, y_j\} \in T$ is the j^{th} trust sample, given x_j is a column vector and y_j is a scalar.

One idea to compute eq. (1) in dZTA is to compute the trust locally (local learning) at each WFH device applying the stochastic descent (gradient), followed by a network-wide global aggregated learning. Given finite step size, similar to [24], the local trust vector denoted by g_n can be updated in each iteration by

$$g_n^{t+1} = g_n^t - \delta \frac{((\nabla F_j(g_n^t) - \nabla F_j(G_T)) + \nabla F(G_T))}{\text{No. of Iterations}}, \quad (2)$$

$$\text{where } F_j(g) = \frac{(x_j^\top g - y_j)^2}{2}.$$

In summary, any WFH device in dZTA setup follows the following steps: *updating* the local learning (g_n^{t+1}), and *uploading* the local trust update ($g_n, \nabla F_j(G_T)$) to the central server. Thereafter, the computation of the global trust update, i.e. ($G_T, \nabla F(G_T)$) happens at the global server. However, in our proposed dZTA system, the main global server in FL is ablated with Blockchain mechanism, which makes the computation of the global trust update possible autonomously within each WFH device (explained next in the subsection below, see Algorithm 1 and Algorithm 2.)

B. Ablating FL with Blockchain

We ablate FL with blockchain, to design the proposed dZTA framework for sharing local trust model updates and experiences from all remote working devices using a distributed ledger, where the blocks in the Blockchain are generated at the rate of γ blocks per second, carries local model updates (of authentication, authorization, reward, etc.) and the block validation is handled by using M data miners.

In particular, the local update block in the ledger consists of two different parts: *header* and *body*. The *header* block has (i) *the pointer to earlier block*; (ii) *block generation rate* γ (blocks per seconds), and (iii) *outcome of the consensus mechanism*. The *body* part holds (i) computing delay in the WFH device, and (ii) learning update of the local trust model ($G_n, \nabla F_j(G_T)$).

Data miner maintains a separate block for storing local trust updates from all the associated WFH devices and other miners. A block will become a new block in the dZTA system if and only if the data miner produces the hash successfully. This will then be followed immediately by a ‘broadcasting of the generated (new) block’ in the entire mining network,

which (i) responds to all other miners, (ii) halts their ongoing computation, and (iii) adds the block into the ledgers.

Nevertheless, there is a probability that before receiving the broadcasted block from the winning miner, other runner-up data miners generate their own blocks, mandating undesirable updates in the ledger of their nearby data miners (with the lately created block)³. Such undesirable impacts, called *forking* [25], frequently occurs in a large distributed network, and therefore need to consider cautiously.

We observe that, γ is adjustable by tuning the difficulty of *proof-of-work*, i.e., the block generation rate γ in the network is often determined by the complexity of the underlying *proof-of-work* procedure. The higher the expected block generation rate γ^* for example, the larger is the target value for hashing. Here, the occurrence of unwanted forking is based on local computing and communication network parameters: γ and the transmission delay. Moreover, the occurrence of forking (and its probability) decreases with a decrease in γ and communication delay/impairments. It is worthy to note that forking can be ameliorated by cautiously tuning the computing (i.e., γ , the block generation rate) and communication delay of the system.

C. Learning Reputation in dZTA

Our mining reputation idea for dZTA states that “*data miners are proportionally rewarded based on their contributed mining efficiency over time t* ”, along the lines of the Blockchain’s traditional approach. The efficiency of a data miner depends on the assembling of the valuable authentication and trust samples from its interlinked WFH devices.

In addition, the implanted Blockchain into FL has the required capability of maintaining a reward program for the promotion and rating of WFH devices as anticipated. The authentication, authorization and trust reward to the device is always proportional to the size of their local samples from the WFH device and can be provided by the system’s associated miner. As a result, with dZTA, We can offer offers a reputation for authentication, authorization and trust samples to the devices, e.g., (*trust reward*). Also, it supplies rewards for the block verification process (*mining reward* for the data miners). Figure 3 shows the flows of authentication, authorization, trust and mining rewards in the dZTA system.

VI. PERFORMANCE EVALUATION

With relevant insights from the above analysis and assumptions, we design guidelines and sketch a high-level skeleton of the two core algorithms. See Algorithm 1 and Algorithm 2 for details.

ALGORITHM 1: Local and Global learning mechanisms at the network node or WFH device. It develops an intelligent agent for the situational awareness and continuous trust evaluation; and

ALGORITHM 2: Algorithm for block mining process at the miners in the cloud. This handles granular access control for dynamic authentication and authorization.

³Other contending runner-up miners are likely to push their generated blocks in the neighbouring miners’ ledgers before obtaining the broadcasted block from the winning miner.

Algorithm 1 dZTA Learning in the WFH device

```

1: procedure LOCALLEARNING
2:   loop No. of iterations
3:     WFH device  $n$  computes new trust  $g_n^{t+1}$ 
4:   end loop
5: end procedure
6: procedure UPLOADTRUST
7:   Run association:  $n$  with  $m$ 
8:   Upload  $n^{th}$  device (updated trust model)
9: end procedure
10: procedure GLOBALLEARNING
11:   while ( $|G_T^{t+1} - G_T^t| > threshold$ ) do
12:     Update  $G_T$ 
13:   end while
14: end procedure
15: procedure DOWNLOADGLOBALTRUST
16:    $n^{th}$  device downloading the block from data miner  $m$ 
17:   device  $n$  uploads local updates and computation time
18: end procedure

```

Algorithm 2 dZTA Mining procedure at data miners

```

1: procedure GENERATEBLOCK
2:   while (receives new) or ( $target == hash$ ) do
3:     Run proof-of-work
4:     if ( $hash == target$ ) then
5:       New block(block): TRANSMITNEWBLOCK
6:       if receives newly generated block then
7:         Check(forking): Send Acks
8:       end if
9:     end if
10:   end while
11: end procedure
12: procedure VERIFICATION
13:    $m^{th}$  data miner broadcasting the model update
14:   while ( $size, timeout, verify$ ) do
15:     Transform local learning into a new block
16:     Update block size (block size + local update)
17:   end while
18: end procedure
19: procedure TRANSMITNEWBLOCK
20:   Broadcasting block
21:   Avoid forking: Check timeouts
22: end procedure

```

A. Results and Discussion

Using Algorithm 1 and Algorithm 2, in Figure 4 and Figure 5, we have shown the overall impact on dZTA operationalization due to the increase in γ and increase in link errors respectively (it performs as anticipated and discussed in Section V-B). More importantly, Figure 4 demonstrates that for a given quality of a link⁴, the trust computation delay is controllable by adequately tuning the block generation rate γ . This observation reflects the effectiveness and flexibility of the fine-grained access control

⁴Wireless link quality are generally measured in terms of signal to noise ratio, which is the measure in decibel (dB) of the difference between the received signal and the errors in the link.

in dZTA using blocks. Figure 5 provides us with an important finding that in erroneous wireless conditions, forking probability increases abruptly and adversely impacts the performance of dZTA operationalization, which requires further consideration. We have observed and quantified the trend of the occurrence of forking, which decreases with a decrease in the block generation rate and wireless impairments.

Next, we perform a preliminary analysis to compare the global accuracy of the learning process using fZTA (FL-based ZTA) and our dZTA operationalization framework. Figure 6 shows that the global model accuracy of fZTA can not withstand well in the presence of malicious devices. Note that we have emulated Gaussian noise into their learning of the local models. However, our proposed dZTA operationalization is significantly robust against such untrusted/malicious nodes, and its accuracy decreases (almost) linearly with an increase in the number of malicious devices. This is in sharp contrast to that of fZTA— please compare the red and green curves in Fig. 6. More importantly, we found that fZTA is highly susceptible to malicious data from untrusted nodes and their activity and, therefore, may not practicable directly for deploying the ZTA in enterprise WFH networks. Observe in Fig. 7 that fZTA suffers heavily even in the presence of 3 malicious nodes when the total population is 20 devices (8 WFH devices). This provides us with confidence that our design of dZTA is essential for deploying ZTA in the proposed distributed WFH environment and it can substantially increases the accuracy of learning processes. Straightforward fZTA using Google's FL appears to perform well only in an ideal network settings without malicious devices⁵.

Furthermore, we apply *Earth mover's distance* to measure the quality of learning data from the malicious devices. The Earth mover's distance is defined by the *percentage probability distributions of the trust model parameters obtained from a tagged device compared to its global distribution* of the whole population of users and 8 WFH devices. Figures 8 and 9 illustrate the maximum perceived global accuracy of the trust computation concerning a different number of malicious WFH devices and Earth mover's distances. Observe

⁵The problem of runtime verification of the proposed dZTA without sharing a global clock requires further investigations along the lines of [35] and will be investigated in future.

Figure 4: dZTA trust computation delay with an increase in γ and three communication costs.

Figure 5: Forking probabilities with increasing channel errors and different γ .

Figure 6: Comparison of global accuracy (fZTA vs dZTA) in the presence of malicious users and 20 devices (8 WFH devices).

Figure 7: Comparison of global accuracy (fZTA vs. dZTA) in the presence of malicious users and 20 devices (8 WFH devices).

Figure 8: Comparison of the maximum accuracy (fZTA vs dZTA) with the increasing number of malicious WFH devices and percentage of Earth mover's distances.

in Fig. 8 that three variables are impacting the accuracy of the learning process: Earth mover's distance, number of malicious nodes, and the population of network and WFH devices. Also, the unreliable and untrusted malicious devices and their training parameters often negatively impact the accuracy. However, for a given population in enterprise network with WFH devices, we can clearly observe that dZTA outperforms fZTA with a finite number of malicious nodes (and when it increases). In particular, the impact on the accuracy of dZTA operationalization with the increase in the number of malicious nodes and Earth mover's distance is far less than that observed on fZTA (see Fig. 8). fZTA performs better than dZTA when the number of malicious nodes is small (see Fig. 9). Overall, for realistic size enterprise WFH network, our proposed dZTA framework implements small trusted zones with high accuracy and significantly outperforms fZTA.

B. Future Directions

One of the important performance indicators for the proposed dZTA is the adaptive selection of optimal block size (i.e., granularity) for maintaining the distributed trust and distributed identities. The block size impacts the overall capabilities of the remote WFH devices and the communication cost (uploading/downloading delay). A detailed study of this effect, accounting for forking and flexible block sizing to understand

(and optimize) their tradeoffs, requires further investigations and will be one of our main focus in future research.

Furthermore, from our proof-of-concept implementation, we found a few important insights towards other measures for evaluating the dZTA. For example, in Figure 4, it appears that there exists an optimal tuple, $(\gamma^*, \text{delay}^*(\text{dZTA}))$, (in the sense that for dZTA to be highly responsive) for minimizing the authentication delay. Similarly, to minimize forking due to errors in the link (recall Figure 5), one may utilize a sufficiently large number of retries for a block until success. However, on the other hand, Figure 5 demonstrated that such repeated attempts and the overall delay might jointly undermine the benefits of dZTA after a certain threshold.

In fact, the condition $(\text{delay}(\text{dZTA}) > \text{threshold})$ can be observed to be context-aware and possibly guided by the tolerable authentication/authorization delay threshold of the WFH device or user application. Other important measures to be considered while developing the components of the proposed dZTA framework are the metrics of the respective domains, such as reputation-based allocation (e.g., trust reward, mining reward as mentioned earlier), the accuracy of the FL process, confidence/trust level, modelling time-varying adversaries, defenses of property and membership inference, and its scalability and flexibility, among others.

Figure 9: Different view of Figure 9: Comparison of the maximum accuracy (fZTA vs dZTA) with the increasing number of malicious WFH devices. This view is showing the area where fZTA performs better than dZTA when the number of malicious WFH device is small.

VII. CONCLUSION

We have categorized ZTA operationalizations and developed a comprehensive framework for federated and decentralized operationalization of ZTA. To meet ZTA's design goals as outlined by NIST, we focus on continuous authentication and authorization, and small implicit trust zones while preserving availability. We account for minimizing temporal delays in authentication techniques to reduce uncertainty. We also enforce the least privileges required to complete the action in the request by creating access rules and making them as granular as feasible. The proposed dZTA framework efficiently facilitates distributed trust computation over WFH devices and remote networks. The developed procedures shed light on granular access control mechanisms for dZTA by employing Blockchain and FL techniques. Expected outcomes of this initiative include advancements towards full-fledged dZTA over enterprise WFH networks with continuous trust evaluation and reward mechanisms.

REFERENCES

- [1] E. Bertino, "Privacy in the era of 5G, IoT, big data and machine learning," in *2020 Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)*, 2020, pp. 134–137.
- [2] B. Chen, S. Qiao, J. Zhao, D. Liu, X. Shi, M. Lyu, H. Chen, H. Lu, and Y. Zhai, "A security awareness and protection system for 5G smart healthcare based on zero-trust architecture," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10 248–10 263, 2021.
- [3] F. Malecki, "Overcoming the security risks of remote working," *Computer Fraud & Security*, vol. 2020, no. 7, pp. 10–12, 2020.
- [4] B. Bonakdarpour, J. V. Deshmukh, and M. Pajic, "Opportunities and challenges in monitoring cyber-physical systems security," in *International Symposium on Leveraging Applications of Formal Methods*. Springer, 2018, pp. 9–18.
- [5] M. Campbell, "Beyond zero trust: Trust is a vulnerability," *Computer*, vol. 53, no. 10, pp. 110–113, 2020.
- [6] E. Bertino, "Zero trust architecture: Does it help?" *IEEE Security & Privacy*, vol. 19, no. 05, pp. 95–96, 2021.
- [7] B. Embrey, "The top three factors driving zero trust adoption," *Computer Fraud & Security*, vol. 2020, no. 9, pp. 13–15, Sep. 2020.
- [8] J. Kindervag et al., "Build security into your network's dna: The zero trust network architecture," *Forrester Research Inc.*, pp. 1–26, 2010.
- [9] R. Ward and B. Beyer, "Beyondcorp: a new approach to enterprise security," *magazine of USENIX & SAGE*, vol. 39, no. 6, pp. 6–11, 2014.
- [10] D. CeArley et al., "Top 10 strategic technology trends for 2018," *The Top*, vol. 10, pp. 1–246, 2016.
- [11] C. Cunningham and Z. T. P. A. Emerging, "The zero trust extended (ztx) ecosystem," *Forrester*, Cambridge, MA, 2018.
- [12] C. Hines, "DP, ZTNA, and CARTA: Making sense of the zero trust security buzz," *ZScaler Blog*, 2019.
- [13] S. Rose et al., "Zero Trust Architecture," National Institute of Standards and Technology, Tech. Rep., Aug. 2020.
- [14] E. Bertino and K. Brancik, "Services for zero trust architectures-a research roadmap," in *2021 IEEE International Conference on Web Services (ICWS)*. IEEE, 2021, pp. 14–20.
- [15] B. Yu et al., "IoTChain: Establishing Trust in the Internet of Things Ecosystem Using Blockchain," *IEEE Cloud Computing*, vol. 5, no. 4, pp. 12–23, Jul. 2018.
- [16] M. A. Azad, S. Bag, F. Hao, and A. Shalaginov, "Decentralized self-enforcing trust management system for social internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 2690–2703, 2020.
- [17] O. Hasan, L. Brunie, E. Bertino, and N. Shang, "A decentralized privacy preserving reputation protocol for the malicious adversarial model," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 949–962, 2013.
- [18] S. Rose, "Planning for a zero trust architecture: A starting guide for administrators (draft)," National Institute of Standards and Technology, Tech. Rep., 2021.
- [19] B. McMahan et al., "Communication-efficient learning of deep networks from decentralized data," in *Artificial Intelligence and Statistics*. PMLR, 2017, pp. 1273–1282.
- [20] K. Hatakeyama et al., "Zero trust federation: Sharing context under user control towards zero trust in identity federation," in *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE, 2021, pp. 514–519.
- [21] P. R. Zimmermann, *The official PGP user's guide*. MIT press, 1995.
- [22] N. Li et al., "Context-Aware Trust Management System for IoT Applications with Multiple Domains," in *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*. Dallas, TX, USA: IEEE, Jul. 2019, pp. 1138–1148.
- [23] L. Alevizos et al., "Augmenting zero trust architecture to endpoints using blockchain: A systematic review," *arXiv preprint arXiv:2104.00460*, 2021.
- [24] S. R. Pokhrel and J. Choi, "Federated Learning With Blockchain for Autonomous Vehicles: Analysis and Design Challenges," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4734–4746, Aug. 2020.
- [25] S. R. Pokhrel, "Federated learning meets blockchain at 6G edge: A drone-assisted networking for disaster response," in *Proceedings of the 2nd ACM MobiCom Workshop on Drone Assisted Wireless Communications for 5G and Beyond*. London United Kingdom: ACM, Sep. 2020, pp. 49–54.
- [26] E. Puggioni et al., "Towards Decentralized IoT Updates Delivery Leveraging Blockchain and Zero-Knowledge Proofs," *arXiv:2010.12134 [cs]*, Oct. 2020.
- [27] I. Ahmed et al., "Protection of Sensitive Data in Zero Trust Model," in *Proceedings of the International Conference on Computing Advancements*. Dhaka Bangladesh: ACM, Jan. 2020, pp. 1–5.
- [28] W. Zhang et al., "Dataset-Level Attribute Leakage in Collaborative Learning," *arXiv:2006.07267 [cs, stat]*, Oct. 2020.
- [29] M. Samaniego and R. Deters, "Zero-Trust Hierarchical Management in IoT," in *2018 IEEE International Congress on Internet of Things (ICIOT)*. San Francisco, CA: IEEE, Jul. 2018, pp. 88–95.
- [30] S. R. Pokhrel et al., "Policy-based Bigdata Security and QoS Framework for SDN/IoT: An Analytic Approach," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. Paris, France: IEEE, Apr. 2019, pp. 73–78.
- [31] C. DeCusatis et al., "Identity-based network security for commercial blockchain services," in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*. Las Vegas, NV: IEEE, Jan. 2018, pp. 474–477.
- [32] P. Kairouz et al., "Advances and Open Problems in Federated Learning," *arXiv:1912.04977 [cs, stat]*, Dec. 2019.
- [33] T. Zhu et al., "Differentially private model publishing in cyber physical systems," *Future Generation Computer Systems*, vol. 108, pp. 1297–1306, 2020.
- [34] P. Xiong et al., "Private collaborative filtering under untrusted recommender server," *Future Generation Computer Systems*, 2018.
- [35] R. Ganguly, A. Momtaz, and B. Bonakdarpour, "Distributed runtime verification under partial synchrony," in *24th International Conference on Principles of Distributed Systems (OPODIS 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021.