ZETInChat: Zero Trust Infrastructure with Dynamic Service Deployment via Chatbot in Mesh Networks

Guilherme Nunes Nasseh Barbosa, Diogo Menezes Ferrazani Mattos
LabGen/MídiaCom – PPGEET/TCE/TET
Universidade Federal Flumimense – UFF – Brazil
{*gbarbosa,menezes*}*@midiacom.uff.br*

***Abstract*—The increasing frequency and severity of cyberattacks demand robust security solutions, especially for mission-critical environments. A key challenge is the dynamic configuration of secure services in Cloud Continuum for ad-hoc mesh networks, which traditional methods struggle to address. In this work, we propose ZETIn, a Zero Trust Infrastructure that leverages the power of Generative AI and a chatbot with LLMs to automatically configure services based on natural language inputs. The proposal ensures end-to-end security, simplifies the configuration process, reduces human error, and enhances network resilience and adaptability. Preliminary results demonstrate effective dynamic service configuration, high security standards, and improved network resilience, demonstrating that the integration of Generative AI with Zero Trust principles is a significant step forward in enhancing security and efficiency in mission-critical ad-hoc networks.**

***Index Terms*—Generative Artificial Inteligence, Zero Trust Infrastructure, Large Language Model, Mesh Network**

I. Introduction

In recent years, cyberattacks have significantly increased in frequency and severity. In response to this trend, the *National Institute of Standards and Technology* (NIST) developed the *Zero Trust Architecture* [1], [2]. The Zero Trust Architecture (ZTA) aims to provide a cybersecurity approach that minimizes implicit trust while requires continuous authentication, strict access control, and network micro-segmentation using Software Defined Perimeter (SDP) [3]. A key challenge is automatically creating secure services in Cloud Continuum environments for mission-critical Ad-Hoc Mesh Networks for autonomous systems, which are complex and error-prone to manage. These networks, essential for tasks like environmental monitoring and military operations, need flexible and secure configurations to ensure data resilience and integrity. Integrating wireless devices adds complexity and vulnerability, underscoring the necessity of ZTA for continuous security.

In this work, we propose and demonstrate the ZETInChat, a Zero Trust platform that leverages Generative AI to automatically configure the network and services from natural language command descriptions. The system enables an interactive chatbot, integrated with large language models (LLMs), to receive commands and queries in natural language from operators and, based on these inputs, dynamically configure the necessary services on the mesh network. It includes applying security policies, allocating resources, and managing secure device communications. The platform uses advanced encryption and authentication techniques to ensure end-to-end security from the cloud to edge devices. This approach simplifies the configuration process and enhances the network's resilience and adaptability to different operational scenarios, improving incident response and protection against threats. The proposal aligns with the needs of autonomous systems, providing robust security and dynamic adaptability required for their operations.

Compared to existing solutions in the literature, our proposal offers a significant advantage by combining Zero Trust Architecture with Generative AI and mission-critical mesh networks. While many current approaches focus on static configurations or require manual adjustment intervention, our solution fully automates the process, reducing the risk of human error and speeding up the implementation of security measures. Additionally, integrating a chatbot with LLMs provides a more intuitive and accessible interface for operators, making network management easier even for those with less technical knowledge. Traditional solutions often fail to provide the flexibility and dynamics needed for mission-critical environments. Conversely, our proposal meets these demands, offering robust and adaptive security that adjusts to real-time needs.

II. Related Work

Previous works utilize Zero Trust Architecture concepts to tackle specific issues in companies and large organizations. Our proposal stands out because it integrates a Zero Trust platform with generative AI capabilities to configure services dynamically in mission-critical mesh networks.

Al-Hammuri *et al.* propose a Zero-trust-based scoring system to prevent medical errors in cloud-based healthcare information systems using machine learning and microservice-based authentication [4]. Our approach focuses on automatic network configuration and management based on uses OpenZiti and Ollama to implement a secure and flexible solution that processes natural language commands, facilitating automated and secure network and service configuration.

Kroculick focuses on theoretical development methodologies for Zero Trust architectures [5]. Our proposal, however, takes a more practical approach. We incorporate an LLM-based chatbot that allows fewer technical operators to dynamically configure complex networks, making the solution more accessible and applicable in real-world scenarios.

This work was carried out with resources from the CNPq, FAPERJ, and CAPES.

Tanimoto *et al.* address the scalability of software-defined perimeters (SDP) in diverse organizations by proposing models such as hierarchical and bridge [6]. In contrast, our project combines SDP with AI capabilities to not only scale but also adapt network configuration according to changes in the operational environment automatically.

Chandramouli and Butcher focus on access control in cloud-native applications using a service mesh and proxy infrastructure [7]. In contrast, our solution integrates automated service detection and configuration through natural language interactions, enhancing usability and security in Mesh networks.

While other works propose specific solutions for trust evaluation and network management [8], [9], our approach is more comprehensive. We offer a holistic solution that not only evaluates trust but also dynamically configures and manages the network securely. By using OpenZiti and generative AI, we ensure a resilient and adaptable infrastructure for various mission-critical scenarios.

III. Zero Trust Architecture and Mesh Network

The proposed ZETIn platform relies on open-source solutions for deploying the Zero Trust Architecture (ZTA), the Mesh Network, and the Artificial Intelligence environment. OpenZiti[1] provides secure overlay networks, implementing ZTA with the Ziti Controller managing the Software Defined Perimeter (SDP) and Edge Routers acting as SDP Gateways. OpenZiti ensures strict access control, network traffic optimization, and scalability, supporting organizational expansion and robust IoT security [10], [11]. The mesh network uses B.A.T.M.A.N.[2] (Better Approach to Mobile Ad-hoc Networking), a proactive routing protocol for Wireless Ad-hoc Mesh Networks, including MANETs. B.A.T.M.A.N. maintains information about accessible nodes, determining the best single-hop neighbor for each destination and facilitating efficient multi-hop routing[3]. The AI implementation leverages Open WebUI[4] and Ollama[5], creating a user-friendly and secure environment. Open WebUI, a self-hosted WebUI, supports various LLM runners, including Ollama and OpenAI-compatible APIs, ensuring data privacy and customization. Ollama enhances AI capabilities by processing natural language commands for dynamic network configurations. This integration provides a reliable and adaptable solution for managing network services, meeting the needs of modern, mission-critical environments.

IV. The ZETInChat Proposal

Figure 1 illustrates the ZETin architecture for a robust, secure, and dynamic solution for configuring and managing mesh networks in mission-critical environments, utilizing Generative AI and Zero Trust principles. The architecture comprises two primary planes: the `Network Control Plane` and the `ZTA Plane`. Additionally, a `Forwarding Plane`, not shown in the figure, handles the actual data transmission.

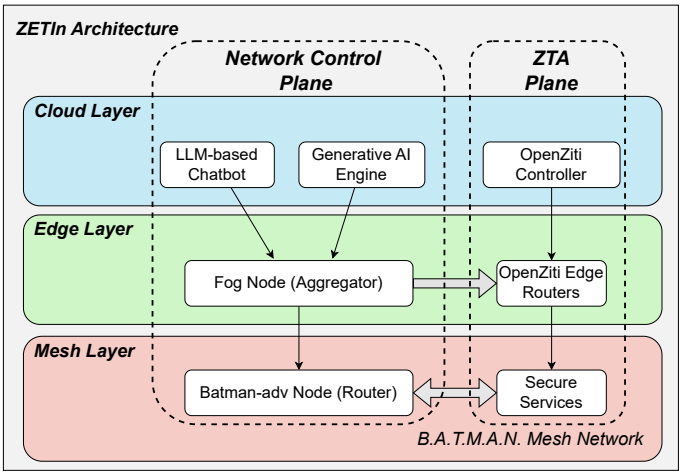

Fig. 1. The ZETIn architecture for managing mesh networks in mission-critical environments using Generative AI and Zero Trust. The Cloud layer includes an OpenZiti Controller, a Generative AI Engine, and an LLM-based chatbot for user interaction. The Edge layer has fog nodes and OpenZiti routers enforcing security. The mesh network uses B.A.T.M.A.N. for mesh communication, secured by Zero Trust. Users interact via a chatbot, which configures the network and services dynamically, ensuring encrypted communication and continuous security.

At the Cloud layer of the `Network Control Plane`, the OpenZiti Controller is responsible for managing the Zero Trust security policy, ensuring continuous authentication and authorization of all access. The Generative AI Engine, a key component in this layer, processes natural language commands, enabling the dynamic configuration of network services. Ollama realizes the Generative AI Engine[6]. This dynamic configuration capability significantly enhances the network's adaptability and responsiveness. Additionally, an LLM-based chatbot provides a user-friendly interface for interaction, allowing users to issue commands and queries in natural language, which are then interpreted and executed by the AI engine. The LLM-based chatbot is deployed with OpenWebUI.

The Edge layer, a crucial part of the `Network Control Plane`, includes fog nodes that aggregate data from IoT devices and serve as intermediaries between the Cloud and the network's edge. These fog nodes provide local processing capabilities and enhance the network's scalability and responsiveness. Equally important in this layer are the OpenZiti edge routers, which play a key role in network security. They enforce security policies, ensuring that all communication is encrypted and that only authenticated and authorized devices can access network resources. This robust security measure significantly enhances the network's security. In the `ZTA Plane`, the mesh network leverages Batman-adv for efficient communication between devices.

[1]https://openziti.io/

[2]https://www.open-mesh.org/projects/open-mesh/wiki

[3]Our proposal is agnostic to the B.A.T.M.A.N. version used, as long as IP connectivity between network nodes is ensured.

[4]https://openwebui.com/

[5]https://ollama.com/

[6]The proposal aims to leverage models such as LLaMA 3.1 and Falcon.

Security within the mesh network is guaranteed by the Zero Trust architecture principles, which are enforced by the policies set by the OpenZiti Controller and executed by the edge routers. Users interact with the system through the chatbot interface, which runs in the Cloud and interprets natural language commands to dynamically configure the network and its services. This interaction allows for real-time adjustments and management of the network, enhancing its adaptability and robustness. The continuous application of security policies and the encryption of communication channels ensure that the network remains secure against unauthorized access and potential threats.

The architecture's design, encompassing the `Network Control Plane`, `ZTA Plane`, and the underlying `Forwarding Plane`, provides a comprehensive and robust solution for managing mission-critical Mesh networks. It combines the strengths of Generative AI and Zero Trust principles to deliver a secure, scalable, and highly responsive network environment, instilling confidence in its ability to meet the demands of mission-critical environments.s in mission-critical environments, utilizing Generative AI and Zero Trust principles. In the Cloud layer, the OpenZiti Controller manages the Zero Trust security policy, the Generative AI Engine processes natural language commands, and an LLM-based chatbot interacts with users. The Edge layer includes fog nodes that aggregate data from IoT devices and act as intermediaries, while OpenZiti edge routers enforce security policies. The Mesh network uses Batman-adv for efficient communication between IoT devices, with security ensured by the Zero Trust architecture. Users interact with the system through a chatbot interface that interprets natural language commands and dynamically configures the network and services. Security is ensured through encrypted communication channels and policies' enforcements.

V. Demonstration

The demonstration will involve a mesh network composed of four Raspberry Pi devices, each acting as a node within the network. These devices will utilize the B.A.T.M.A.N. protocol for communication and routing. A gateway will connect the mesh network to the Cloud, where the OpenZiti Controller, Generative AI Engine, and LLM-based chatbot modules are deployed. The demonstration will showcase the dynamic configuration of network services through natural language commands processed by the chatbot, highlighting real-time adjustments and security policy enforcement. This setup will illustrate the seamless integration between the edge devices and the Cloud, ensuring secure, adaptable, and efficient management of the Mesh network in a mission-critical environment. The prototype is currently at Technology Readiness Level (TRL) 4, progressing towards TRL 5[7], as it has been validated in a lab environment and is now ready to be tested for performance and reliability in an operational setting.

[7]The demonstration will involve field testing to validate reliability, and expanding experiments with more nodes to assess scalability and performance in complex environments.

VI. Conclusion

The proposed ZETInChat platform integrates Generative AI with Zero Trust Architecture to address the complex challenges of dynamic service configuration in mission-critical mesh networks for autonomous systems. By leveraging OpenZiti and Ollama, the system ensures end-to-end security, simplifies configuration processes, reduces human error, and enhances network resilience and adaptability. Compared to existing solutions, our approach stands out by offering automated, real-time configuration via a user-friendly chatbot interface, ensuring continuous security and efficient management of network services. The demonstrated effectiveness in preliminary results, TLR4, underscores the potential of combining Generative AI and Zero Trust principles to significantly enhance the security and efficiency of mission-critical ad-hoc mesh networks. The proposal is in line to be adopted by autonomous systems, providing them with robust security and dynamic adaptability required for autonomous operations. Future work focuses on improving scalability for larger network environments, integrating advanced threat detection mechanisms using machine learning. Additionally, we envision analyzing the potential vulnerabilities of using LLMs, such as adversarial attacks and natural language processing errors.

References

[1] N. Sheikh, M. Pawar, and V. Lawrence, "Zero trust using network micro segmentation," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2021, pp. 1–6.

[2] A. Zivi and C. Doerr, "Adding zero trust in byod environments through network inspection," in *2022 IEEE Conference on Communications and Network Security (CNS)*, 2022, pp. 1–6.

[3] N. F. Syed, S. W. Shah, A. Shaghaghi, A. Anwar, Z. Baig, and R. Doss, "Zero trust architecture (zta): A comprehensive survey," *IEEE Access*, vol. 10, pp. 57 143–57 179, 2022.

[4] K. Al-hammuri, F. Gebali, and A. Kanan, "Ztcloudguard: Zero trust context-aware access management framework to avoid medical errors in the era of generative ai and cloud-based health information ecosystems," *AI*, vol. 5, no. 3, pp. 1111–1131, 2024. [Online]. Available: https://www.mdpi.com/2673-2688/5/3/55

[5] J. B. Kroculick, "Zero trust decision analysis for next generation networks," in *Disruptive Technologies in Information Sciences VIII*, vol. 13058. SPIE, 2024, pp. 278–286.

[6] S. Tanimoto, P. Yangchen, H. Sato, and A. Kanai, "Suitable scalability management model for software-defined perimeter based on zero-trust model," *International Journal of Service and Knowledge Management*, vol. 7, no. 1, 2023.

[7] R. Chandramouli and Z. Butcher, "A zero trust architecture model for access control in cloud-native applications in multi-cloud environments," National Institute of Standards and Technology, Tech. Rep., 2023.

[8] S. A. Khowaja, L. Nkenyereye, P. Khowaja, K. Dev, and D. Niyato, "Slip: Self-supervised learning based model inversion and poisoning detection-based zero-trust systems for vehicular networks," *IEEE Wireless Communications*, vol. 31, no. 2, pp. 50–57, 2024.

[9] R. Alboqmi, S. Jahan, and R. F. Gamble, "A runtime trust evaluation mechanism in the service mesh architecture," in *2023 10th International Conference on Future Internet of Things and Cloud (FiCloud)*, 2023, pp. 242–249.

[10] S. Teerakanok, T. Uehara, and A. Inomata, "Migrating to zero trust architecture: Reviews and challenges," *Security and Communication Networks*, vol. 2021, no. 1, p. 9947347, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2021/9947347

[11] J. J. Diaz Rivera, T. A. Khan, W. Akbar, A. Muhammad, and W.-C. Song, "ZT&T: Secure blockchain-based tokens for service session management in zero trust networks," in *2022 6th Cyber Security in Networking Conference (CSNet)*, 2022, pp. 1–7.