Intelligent monitoring system

Enhancing Network Monitoring with Zabbix, Grafana, and Machine Learning

**Luís António Pinto de Barros**

Polytechnic Institute of Guarda

**Abstract —** During my internship at Securnet, I developed a project focused on building a network laboratory with a DMZ, where I deployed SFTP, email, and web servers. The main objective was to create a controlled environment for monitoring and optimizing network resources.

The project began with the design and configuration of the network, including the setup of a DMZ and essential servers. I then implemented the Zabbix monitoring system, enabling real-time supervision of network resources and proactive issue detection. To enhance monitoring efficiency, I configured hosts to ensure comprehensive visibility across all machines in the lab.

Additionally, I integrated Zabbix with the GLPi asset management system, providing a structured and centralized view of monitored devices. I also connected Zabbix with Grafana, facilitating the creation of customized dashboards for improved data visualization and analysis.

To further enhance monitoring capabilities, I applied machine learning techniques to the data collected by Zabbix, aiming to reduce false positives and filter out irrelevant events. This approach improved the accuracy and efficiency of network monitoring, contributing to more effective resource management.

***Keywords - Network engineering; Zabbix; GLPi; Grafana; React Native; Machine Learning***

1. MOTIVATION

The core motivation behind this project stemmed from the growing need for robust network monitoring, particularly in an era where cyber threats are constantly evolving. Establishing a controlled laboratory with a DMZ was essential for simulating real-world scenarios, allowing for the assessment of both the responsiveness and effectiveness of the deployed tools.

Integrating Zabbix with other platforms such as GLPi and Grafana aimed not only to enhance operational efficiency but also to improve data visualization and analysis. Furthermore, incorporating machine learning techniques represented a forward-thinking approach to event filtering, minimizing noise and maximizing the relevance of captured information.

Ultimately, this project was driven by the goal of building a resilient infrastructure that not only strengthens security but also fosters innovation in the field of cybersecurity.

1. PROJECT OBJECTIVES

This project aims to establish a robust and efficient network monitoring environment, focusing on the integration of multiple tools and innovative techniques. To achieve this goal, the following objectives were defined:

* Configure and test a network laboratory with a DMZ and associated servers.
* Implement and optimize the Zabbix monitoring system for continuous network supervision.

**António Mário Ribeiro Martins**

Polytechnic Institute of Guarda

* Integrate Zabbix with the GLPi asset management system to provide a consolidated view of network devices.
* Establish a connection between Zabbix and Grafana to develop custom monitoring dashboards and advanced visual analysis.
* Develop a mobile and web application using React Native to centralize and simplify network monitoring access.
* Apply machine learning techniques to Zabbix data to filter out irrelevant events and enhance detection accuracy.
* Conduct comprehensive system testing to ensure the effectiveness and efficiency of all implementations.
* Document all configurations and procedures to facilitate future expansions or modifications of the system.

1. STATE OF THE ART

Network monitoring and optimization are essential in the digital era, where security and efficiency are paramount. Understanding the technological landscape is crucial for selecting the most effective tools to meet the project’s requirements.

To begin, an analysis of existing network monitoring solutions was conducted, as summarized in Table 1.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Plataforma**  **Características** | Zabbix | Nagios | PRTG | SolarWind s | WireShark | Splunk | Graylog |
| Monitorização em  tempo real | ✔ |  | ✔ | ✔ | ✔ |  |  |
| Integração | ✔ | ✔ | ✔ |  |  | ✔ | ✔ |
| Alertas configuráveis | ✔ | ✔ |  |  |  |  |  |
| Web interface | ✔ | ✔ | ✔ | ✔ |  | ✔ | ✔ |
| Monitorização de  aplicações | ✔ |  |  | ✔ |  |  |  |
| Análise de  pacotes |  |  |  |  | ✔ |  |  |
| Centralização de  logs |  |  |  | ✔ |  | ✔ | ✔ |

*Table 1- Comparison of Features by Software*

* + 1. ***Existing Network Monitoring Solutions***

Among the available options, Zabbix, Nagios, and PRTG were selected for a deeper analysis. Zabbix was chosen as the primary focus of this project, while Nagios and PRTG were included due to their widespread adoption in the industry.

* + 1. ***Critical Analysis***
* **Zabbix** is a highly flexible and extensible end-to-end network monitoring solution.
* **Nagios**, while robust, often requires more manual configuration and customization.
* **PRTG**, on the other hand, is known for its user-friendly interface but may have limitations in scalability.

A comparative assessment of their functionalities is presented in Table 1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Bibliotecas**  **Caractertísticas** | TensorFlow | Scikit-learn | Keras | Pytorch |
| Open source | ✔ | ✔ | ✔ | ✔ |
| Suporte para redes neuronais | ✔ |  | ✔ | ✔ |
| Flexibilidade e modularidade | ✔ | ✔ | ✔ |  |
| Algoritmos predefinidos |  | ✔ |  |  |
| Computação auto diferenciável | ✔ |  |  | ✔ |

* + 1. ***Existing Machine Learning Libraries***

For machine learning integration, **TensorFlow** and **Scikit-learn** were selected as the primary libraries due to their extensive adoption, computational efficiency, and applicability in network anomaly detection.

* **TensorFlow**, with its highly scalable tensor operations and deep learning capabilities, provides an optimized framework for handling high-dimensional time-series data from network telemetry. Its ability to leverage **GPU acceleration, automatic differentiation**, and **graph computation** makes it ideal for complex anomaly detection models, such as **autoencoders for unsupervised learning** and **recurrent neural networks (RNNs)** for predictive network behavior analysis.
* **Scikit-learn**, on the other hand, offers a more streamlined and lightweight machine learning approach, excelling in traditional **supervised and unsupervised learning algorithms**, such as **Random Forests, K-Means clustering, and Support Vector Machines (SVMs)**. Its efficiency in **feature engineering, dimensionality reduction (PCA, t-SNE), and statistical modeling** makes it particularly useful for classifying network events and optimizing alert thresholds.

A comparative evaluation of their **computational performance, scalability, ease of integration with Zabbix, and real-time inference capabilities** is provided in Table 2.

TOPOLOGIA DE REDE

Neste trabalho, configurou-se uma rede simulando a complexidade encontrada em infraestruturas empresariais reais, englobando não só a instalação de dispositivos, mas também a construção de um ecossistema digital análogo ao usado por organizações modernas.

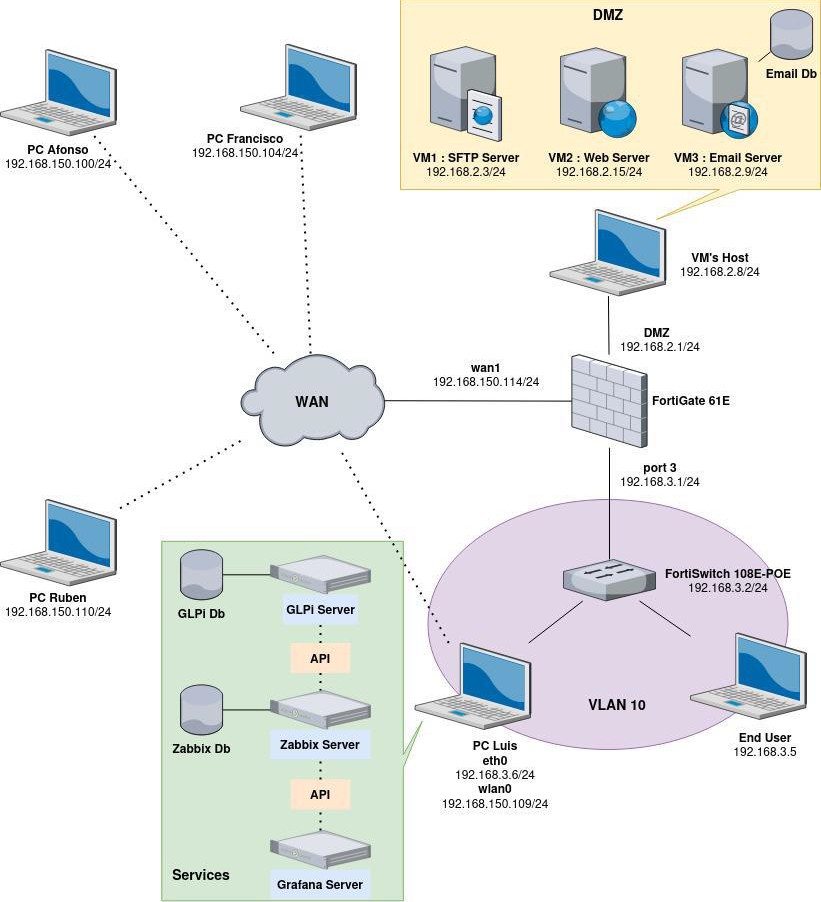
* 1. *Configuração da rede*

O laboratório envolveu a implementação de servidores essenciais: e-mail, web e SFTP. Além disso, foi dada atenção à instalação de uma *firewall* e de um *switch,* componentes cruciais para a conexão e segurança na infraestrutura.

* 1. *Implementação da DMZ*

Dentro das medidas de segurança, enfatizou-se a construção de uma DMZ. Esta camada serve como uma barreira, colocando recursos como servidores web e e-mail em uma área intermediária entre a rede interna e externa, blindando assim ativos sensíveis de ameaças externas.

A arquitetura de rede implementada pode ser visualizada na Figura subsequente, Figura 1.



*Figura 1 - Arquitetura de rede*

Importante ressaltar que, na configuração da firewall, as regras foram estritamente definidas para permitir apenas os protocolos essenciais à comunicação interna da rede e ao acesso á internet. Esta abordagem restritiva assegurou que apenas os protocolos vitais para a operacionalidade dos servidores, comunicação dentro da rede e acesso seguro á internet estivessem autorizados, reduzindo assim possíveis vetores de ataque e garantindo a integridade da rede.

1. zabbix, grafana e glpi

No atual cenário digital, a monitorização e gestão eficaz da infraestrutura de rede é fundamental para garantir a continuidade dos negócios, a segurança e a otimização dos recursos. Neste contexto, a implementação do Zabbix, em conjunto com o Grafana e o GLPi, oferece uma solução integrada para a monitorização, visualização e gestão de incidentes na rede.

1. *Implementação do Zabbix*

O Zabbix foi adotado como a principal ferramenta de monitorização da infraestrutura de rede, dada a sua versatilidade e capacidade de monitorizar uma vasta gama de métricas. Desde a saúde do hardware, passando pelo tráfego da rede até o desempenho de aplicações, o Zabbix proporciona uma visão abrangente do estado operacional da infraestrutura. As alertas e triggers configuráveis garantem que as equipas sejam notificadas proactivamente sobre potenciais problemas, permitindo uma rápida intervenção.

1. *Integração com o Grafana*

Para melhorar a visualização das métricas recolhidas pelo Zabbix, foi implementada uma integração com o Grafana. Esta ferramenta de visualização *open source* amplia as capacidades do Zabbix, permitindo a criação de painéis de controlo interativos e visualmente atrativos. Assim, os administradores de rede têm à disposição uma interface intuitiva que destila as informações complexas da rede em representações gráficas claras e compreensíveis.

1. *Integração com o GLPi*

O GLPi, uma plataforma de gestão de serviços de TI, foi incorporado ao ecossistema para gerir incidentes e pedidos relacionados à infraestrutura de rede. Com a integração do Zabbix ao GLPi, os alertas e notificações podem ser automaticamente transformadas em tickets no sistema de gestão, facilitando a coordenação e a monitorização de incidentes. Esta integração assegura que os problemas identificados sejam devidamente registados, priorizados e resolvidos de forma eficaz e tempestiva.

1. zabbix, grafana e glpi

Face à complexidade crescente das infraestruturas tecnológicas, surge a imperativa necessidade de centralizar o acesso a ferramentas vitais como Zabbix, Grafana e GLPi. Esta centralização visa otimizar fluxos de trabalho e auxiliar na tomada de decisões bem fundamentadas.

A solução proposta é uma aplicação em React Native, tanto móvel quanto web. Esta aplicação unifica o acesso a estas ferramentas em diferentes plataformas, como web, Android e iOS, oferecendo uma interface coesa e intuitiva que agiliza a interação com cada sistema.

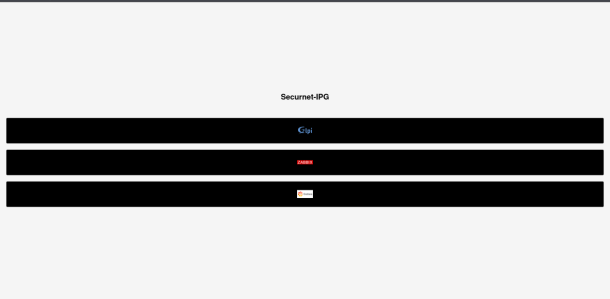
1. *Características do React Native*
   * Desenvolvimento Multi-Plataforma: Permite criar aplicações para múltiplas plataformas a partir de um único código-fonte, otimizando o desenvolvimento e a manutenção.
   * Interface Nativa e Alto Desempenho: Proporciona uma experiência de usuário comparável às aplicações desenvolvidas em linguagens nativas.
   * Reutilização de Componentes: Favorece a consistência e diminui o tempo de desenvolvimento ao reutilizar componentes entre plataformas.
2. *Sobre a aplicação*

A aplicação foi concebida para ser intuitiva, com destaque para botões que direcionam rapidamente os utilizadores para funcionalidades específicas. Esta abordagem simplifica o acesso a serviços essenciais, eliminando a necessidade de memorização de endereços específicos de servidores.

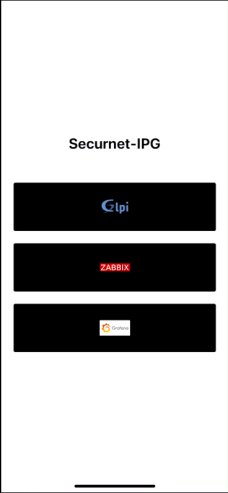
1. *Relevância em ambientes de produção*

Em ambientes que integram múltiplos servidores como Zabbix, GLPi e Grafana, a centralização do acesso é crucial. Com a aplicação React Native, garante-se uma interface singular para todos esses serviços, contando que haja conexão com a rede corporativa, simplificando o processo e aumentando a eficiência operacional.

A aplicação em iOS e web pode é ilustrado nas figuras subsequentes, Figuras 2 e 3.



*Figura 2 - App em web*



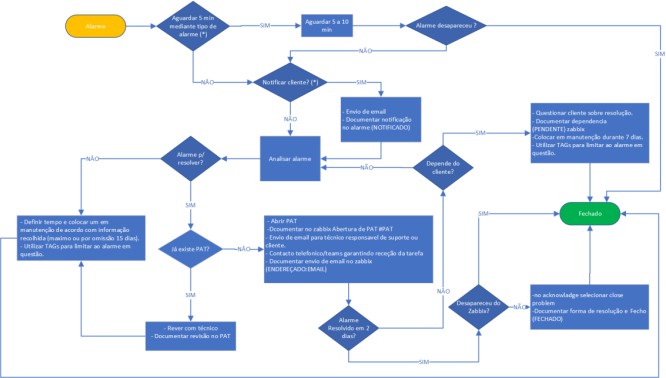
*Figura 3- App em iOS*

1. MACHINE LEARNING

À medida que as empresas modernas são inundadas por uma crescente onda de dados, o imperativo de extrair valor desse mar de informações torna-se mais urgente. O atual cenário digital, com sua constante corrente de dados, detém inúmeras oportunidades. Se interpretado corretamente, esse fluxo pode guiar decisões estratégicas, realçar a eficiência operacional e proporcionar uma vantagem competitiva.

Neste cenário, o *machine learning* destaca-se como uma força transformadora. Com a habilidade de decifrar padrões e irregularidades nos dados, otimizar processos e fundamentar tomadas de decisão, os algoritmos de *machine learning* apresentam-se como aliados valiosos. Considerando a avalanche de eventos que o Zabbix pode gerar, empregar técnicas de *machine learning* para filtrar e priorizar esses eventos torna-se quase obrigatório. Afinal, discernir entre eventos vitais e ruídos em sistemas complexos é um desafio.

A Figura 4 fornece um visual para este processo, ilustrando o fluxograma de ação quando um alarme é disparado no Zabbix.



*Figura 4 - Fluxograma de procedimento após alarme no Zabbix*

Será aprofundada a exploração em torno da escolha e aplicação do algoritmo KMeans para segmentar eventos do Zabbix. Escolhido devido à sua eficácia em tratar dados numéricos contínuos e pela clareza na interpretação dos clusters formados, o KMeans também se destaca pela sua eficiência computacional, um atributo valioso ao lidar com grandes volumes de dados.

KMeans é um pilar no domínio do *clustering*, cuja missão é segmentar um conjunto de dados em K grupos predefinidos, identificando cada cluster pela sua média ou centroide. A operação deste algoritmo abrange a seleção inicial de K centroides, a atribuição de pontos do conjunto de dados ao centroide mais próximo e a calculação contínua dos centroides até a convergência.

Entretanto, como qualquer ferramenta, o KMeans tem as suas particularidades. A sua eficácia pode ser afetada pela seleção inicial de centroides, e seu design pressupõe que os clusters são esféricos e equilibrados. Além disso, é sensível a *outliers*, exigindo cuidados adicionais na preparação dos dados.

Uma ferramenta crucial no arsenal do KMeans é a distância euclidiana, uma métrica que quantifica a separação entre dois pontos em um espaço dimensional. Esta métrica, fundada nos ensinamentos do matemático Euclides, é fundamental para determinar a proximidade e atribuição de pontos aos centroides.

1. *Elbow method*

Primeiramente, foi determinado o número ideal de clusters para segmentar os eventos. Utilizando o *elbow method* com o algoritmo KMeans, constatou-se que o ponto de "cotovelo" indicava um K ótimo de 3 clusters.

1. *Modelagem e processamento de dados*

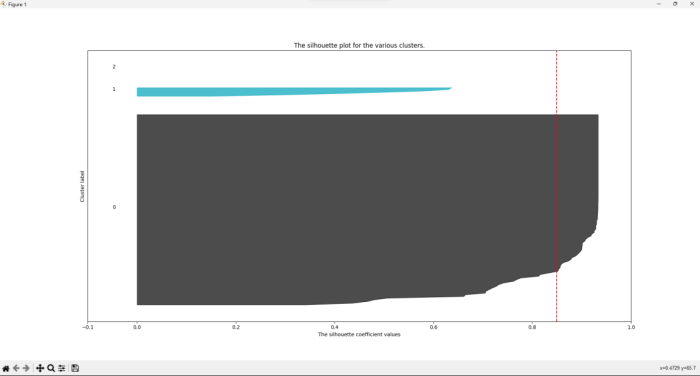
Os dados extraídos do Zabbix foram processados e preparados para serem inseridos no modelo KMeans. Após a aplicação do algoritmo, os eventos foram segmentados em 3 clusters distintos, com base nas suas características intrínsecas.

1. *Avaliação do modelo*

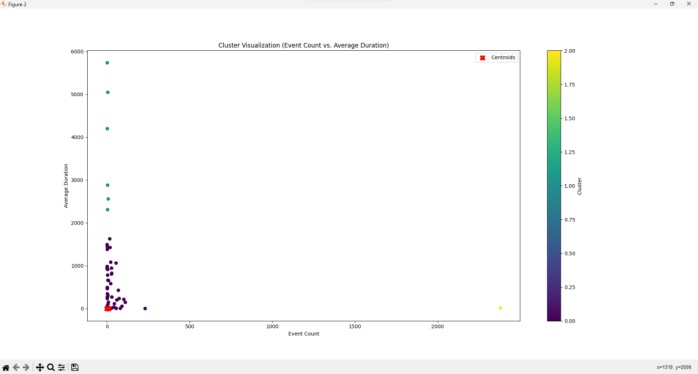
Foi avaliado o desempenho do agrupamento usando o coeficiente de silhueta. Este coeficiente oferece um equilíbrio entre a coesão dos membros de um cluster e a separação entre clusters diferentes. Com um coeficiente de 0.8491, os resultados sugerem uma segmentação significativa dos eventos, com clusters bem definidos.

Para representar a eficácia do agrupamento, geramos gráficos que ilustram o coeficiente de silhueta de cada amostra e a distribuição dos eventos em relação à contagem e duração

média. Isso proporcionou uma visão clara da pertinência de cada evento aos clusters atribuídos. Estes gráficos podem ser visualizados nas figuras seguintes Figura 3 e 4.



*Figura 5 - Coeficiente de silhueta por cluster*



*Figura 6 - Representação gráfica dos clusters*

1. CONCLUSÕES

Combinar monitorização de redes com *Machine Learning* oferece um caminho promissor para a gestão de eventos na era digital. Neste estudo, foi demonstrado um passo nessa direção, e a jornada de inovação continua. Futuros trabalhos podem explorar algoritmos de *clustering* alternativos ou a integração deste modelo diretamente na interface do Zabbix para insights em tempo real.