Applied DevSecOps for Machine Learning Prediction Systems

Abstract

Even though modern IT systems get more and more secure over time, new security vulnerabilities and issues are discovered every day, causing significant economic loss and legal problems. Machine Learning (ML) systems are no exception. To ensure a robust, secure and privacy compliant system, the DevSecOps approach has been implemented on a prediction server for the eVIP (Energy Visualization Integration and Prediction) project. DevSecOps is known to be a methodology that enhances standard DevOps practices and tools by adding a security layer throughout the entire development lifecycle [1].

The eVIP project aims to predict the load curve of electric vehicles (EV) in a semi-private context related to hotels and restaurants in collaboration with OIKEN (regional Distribution System Operator). The eVIP prediction server has been developed and deployed to predict power consumption of pilot hotels for the next 15 minutes. This service enables the eVIP system to automatically set the amps of hotels' charging stations. Using Vehicle to Grid (V2G) protocol, EV batteries can act as a temporary power supply for hotels to optimize peak consumption.

Considering the complexity of machine learning systems, and the fact that it needs several improvement iterations, developers must ensure that no security breach is added by mistake to the code in between releases. To limit risks, we present a DevSecOps pipeline that automates redundant but essential tasks, including unit tests, API tests and security scans. This process automatically deploys updates in production and will identify and block deployment of insecure releases if a vulnerability is found.

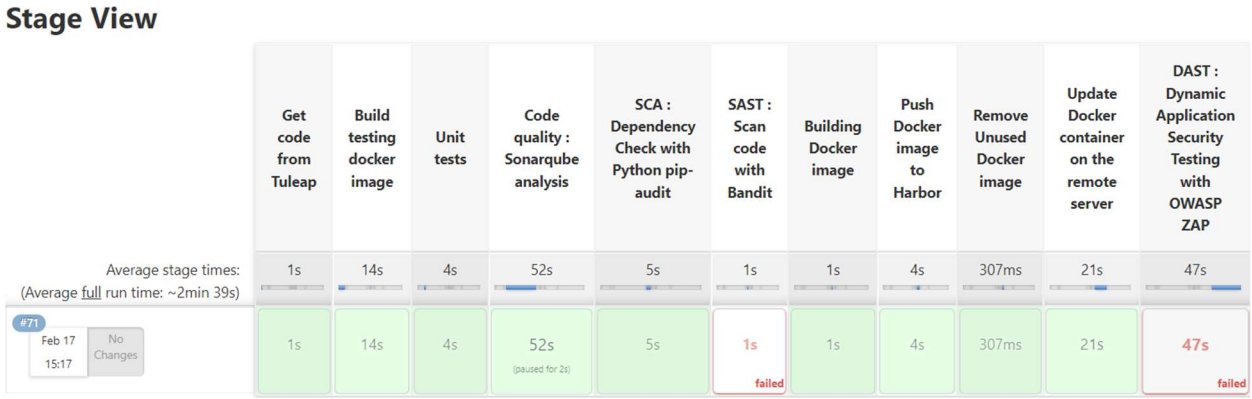

Figure 1 : Jenkins process to ensure a quality DevSecOps process

When implementing such systems, programing frameworks and libraries used for development must be considered. The eVIP prediction server has been developed in Python 3, coupled with Flask library and Docker for deployment. State-of-the-art has been conducted to find the most relevant tools, considering the chosen development framework. Several open-source programs have been compared to find optimal solutions. Compared tool types include software composition analysis (SCA), static and dynamic application security testing (SAST, DAST), container vulnerability Scanner and vulnerability management tools.

Target audience

Data Scientists, IT Managers, DevOps Engineers, Machine Learning Engineers, Developers, IT Security Specialists.

Talk Outline

This talk briefly presents the eVIP project and the eVIP prediction server. Then, you will learn about the stakes and issues linked to software security and personal data processing. We will also summarize the state-of-the-art for selecting the appropriate tools and present an overview of the compared open source softwares and libraries used by the DevSecOps pipeline. Technical aspects of the DevSecOps implementation are presented, including:

- List of DevSecOps tools used and their purposes
- Security enhanced continuous integration pipeline
- Management of vulnerabilities using a vulnerability management tool
- Monitoring and data gathering
- Results and security findings

Below you will find the construction of our presentation:

- Introduction
    - HES – EASILab
    - eVIP project
    - Partners
- Context - Goals
- Prediction server overview
- DevSecOps pipeline for eVIP prediction server
- Advantages and results
- Conclusion

Author biography

David Wannier is a professor at the University of Applied Sciences (UAS) and head of the research team EASILab at the HES-SO Valais-Wallis in Sierre. With 18 years of experience in IT systems integration and software industrialization, he focuses his research activities on new renewable energies (NRE) concerns. He is particularly interested in the development of NRE as well as their potential for evolution through storage in the latest generation batteries.

Jean-Marie Alder is a research assistant at the HES-SO Valais-Wallis in Sierre. He joined the EASILab team in 2020 and has been developing IT solutions to help the energy transition. His experience in software industrialization and DevOps contributed to the implementation of several projects, including the eVIP prediction server.

References

[1] Morales, Jose., Turner, Richard., Miller, Suzanne., Capell, Peter., Place, Patrick., & Shepard, David. (2020). Guide to Implementing DevSecOps for a System of Systems in Highly Regulated Environments (CMU/SEI-2020-TR-002). Retrieved March 01, 2022, from the Software Engineering Institute, Carnegie Mellon University website: http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=638576