Proposal of team software project

Department of Software Engineering

Faculty of Mathematics and Physics, Charles University

Solvers:	Bc. Ivona Oboňová
Study program:	Computer Science - Software and Data Engineering
Project title:	Extraction of GroupRule component into standalone service
Project type:	Company project
Supervisor:	Ing. Pavel Koupil, Ph.D.
Consultants:	Bc. Max Kovykov
Expected start:	2025-04-01
Expected end:	2025-09-01

1 Introduction

At **Avast**, now part of **Gen Digital**, we specialize in cybersecurity software, delivering advanced threat detection powered by AI, along with antivirus protection and privacy solutions such as anti-tracking software, VPNs, and data breach monitoring. Our backend engineering team is at the core of this mission, processing millions of malware samples daily and developing and maintaining internal backend services and APIs that support our malware analysts, security researchers, and other experts in enhancing our products.

One of our core systems, **Appserver**, is a .NET-based monolithic application running on multiple Windows servers. It stores critical data and provides numerous APIs that analysts rely on daily to identify and manage malware threats, making it an essential part of our infrastructure. Among its many components, the **GroupRules** module is responsible for managing string-based and multi-string-based malware definitions and rules. However, as Appserver has grown significantly over time, it has become increasingly difficult to navigate, maintain, and extend with new features.

To address these challenges, our team has begun decoupling the system into a service-based architecture. As part of this initiative, this project focuses on migrating the GroupRules component into a standalone service. This transition will enhance scalability, maintainability, and flexibility. By isolating this component from the monolith, we aim to improve performance and streamline malware definition management, creating a more efficient and future-proof solution.

2 Project description

As mentioned in the introduction, we aim to migrate the **GroupRules** component into a standalone service, decoupling it from the **Appserver**, which has grown increasingly complex, making it difficult to navigate and extend. Additionally, the GroupRules component is embedded within a complex relational database schema. The current implementation uses a PostgreSQL database to store data, which has a largely graph-like structure. As a result, retrieving data in the required format and meeting specific conditions is both complex and inefficient, further complicating maintenance and updates.

To address these issues, we will not only extract GroupRules into its own service, but also evaluate whether a graph database would be a better fit for storing and managing string-based and multi-string-based malware definitions. Given the interconnected nature of the data and its relationships, graph databases appear to be a promising alternative. As part of this project, we will assess their feasibility. If a graph database proves to be the optimal choice, we will identify the best technology and implement GroupRules on top of it. If not, we will continue using the existing relational database while refining its structure where possible.

Rather than rewriting the entire component, which would introduce significant complexity and risk, we will focus on rebuilding only its core functionality as an independent service. This approach allows for gradual integration into the existing infrastructure without disrupting public-facing APIs and UIs. By maintaining unchanged external interfaces, we ensure a smooth and controlled transition, enabling early validation of the new architecture.