

The Manta Flow is a highly automated static analysis platform producing data lineage over its input and representing it in a graph. The platform performs analysis on various technologies and programming languages via specialised scanners. One of the scanners performs analysis of C# code, or rather its compiled alternative Common Intermediate Language. While the scanner was already capable of analysing non-trivial scenarios, it lacked in some aspects that held it up from its broader adoption by customers. The main issues are low support for analysis of real-life scenarios such as web applications or embedded code in other technologies, sub-optimal performance and imprecise lineage output. As a part of this thesis, we improved the precision, scalability and performance of the scanner on multiple levels of abstractions, from analysis of the CIL to modifications of core high-level analysis algorithms. We added support for analysis of the ASP.NET web endpoints and enabled the C# scanner to be used as a service for analysis of embedded code in other technologies. We improved the precision of the resulting lineage for existing scenarios by modifying the core algorithms used throughout the analysis and optimized the analysis process to lift its performance.