Contribution ID: **225** Type: **Sectional reports**

Scalable semantic virtual machine framework for language-agnostic static analysis

Monday, 10 September 2018 14:00 (15 minutes)

The more static program analysis spreads in the industry, the clearer it becomes that its real-world usage requires something more than just optimized algorithms to make execution fast. Distribution of computations to cluster nodes is one of the viable ways for speeding up the process.

The research demonstrates one of the approaches for organizing distributed static analysis - "Semantic Virtual Machines", which is based upon symbolic execution of language-agnostic intermediate codes of programs. These codes belong to the developed language for program representation with semantic objects behind. As objects are fully serializable, they can be saved to and later retrieved from a distributed database. That opens up a possibility for making an analyzer spread to many cluster nodes.
Semantic Hypervisor is the core of the system which acts as a master node for managing runnable virtual machines and also a relay between them and a semantics storage. Main efforts are put into developing of intercommunication mechanisms which can be effective in most of the scenarios, as well as into integration of the approach in existing static analyzer. The study shows the architecture, describes its advantages and disadvantages, suggests solutions for the biggest problems observed during research. The empirical study shows improved performance compared to single node run, and the result is almost linearly scalable due to the high locality of data.

The main result is the created scalable and unified language-agnostic architecture for static analysis with semantics data in a distributed storage behind of it.

The significance of the research is in high achieved performance level alongside with a clean architecture.

Primary author: Mr MENSHCHIKOV, Maxim (St.Petersburg State University)

Presenter: Mr MENSHCHIKOV, Maxim (St.Petersburg State University)

Session Classification: Scientific, industry and business applications in distributed computing systems, education

Track Classification: 4. Scientific, industry and business applications in distributed computing systems