Cache-Based Application Detection in the Cloud Using Machine Learning

Title	Cache-Based Application Detection in the Cloud Using Machine Learning
Publication Type	Conference Paper
Year of Publication	2017
Authors	Gulmezoglu, Berk, Eisenbarth, Thomas, Sunar, Berk
Conference Name	Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security
Publisher	ACM
Conference Location	New York, NY, USA
ISBN Number	978-1-4503-4944-4
Keywords	cross-vm attacks, machine learning, Metrics, prime&probe, privacy, pubcrawl, SVM, threat vectors

Abstract

Cross-VM attacks have emerged as a major threat on commercial clouds. These attacks commonly exploit hardware level leakages on shared physical servers. A co-located machine can readily feel the presence of a co-located instance with a heavy computational load through performance degradation due to contention on shared resources. Shared cache architectures such as the last level cache (LLC) have become a popular leakage source to mount cross-VM attack. By exploiting LLC leakages, researchers have already shown that it is possible to recover fine grain information such as cryptographic keys from popular software libraries. This makes it essential to verify implementations that handle sensitive data across the many versions and numerous target platforms, a task too complicated, error prone and costly to be handled by human beings. Here we propose a machine learning based technique to classify applications according to their cache access profiles. We show that with minimal and simple manual processing steps feature vectors can be used to train models using support vector machines to classify the applications with a high degree of success. The profiling and training steps are completely automated and do not require any inspection or study of the code to be classified. In native execution, we achieve a successful classification rate as high as 98% (L1 cache) and 78$\backslash$% (LLC) over 40 benchmark applications in the Phoronix suite with mild training. In the cross-VM setting on the noisy Amazon EC2 the success rate drops to 60$\backslash$% for a suite of 25 applications. With this initial study we demonstrate that it is possible to train meaningful models to successfully predict applications running in co-located instances.

URL https://dl.acm.org/citation.cfm?doid=3052973.3053036
DOI 10.1145/3052973.3053036
Citation Key gulmezoglu_cache-based_2017

cross-vm attacks machine learning Metrics prime&probe privacy pubcrawl SVM threat vectors