'DAWG' system aims to prevent attacks made possible by Meltdown and Spectre

October 18 2018

Credit: CC0 Public Domain

In January the technology world was rattled by the discovery of Meltdown and Spectre, two major security vulnerabilities in the processors that can be found in virtually every computer on the planet.

Perhaps the most alarming thing about these vulnerabilities is that they didn't stem from normal software bugs or physical CPU problems.

Instead, they arose from the architecture of the processors themselves—that is, the millions of transistors that work together to execute operations.

"These attacks fundamentally changed our understanding of what's trustworthy in a system, and force us to re-examine where we devote security resources," says Ilia Lebedev, a Ph.D. student at MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL). "They've shown that we need to be paying much more attention to the microarchitecture of systems."

Lebedev and his colleagues believe that they've made an important new breakthrough in this field, with an approach that makes it much harder for hackers to cash in on such vulnerabilities. Their method could have immediate applications in cloud computing, especially for fields like medicine and finance that currently limit their cloud-based features because of security concerns.

With Meltdown and Spectre, hackers exploited the fact that operations all take slightly different amounts of time to execute. To use a simplified example, someone who's guessing a PIN might first try combinations "1111" through "9111." If the first eight guesses take the same amount of time, and "9111" takes a nanosecond longer, then that one most likely has at least the "9" right, and the attacker can then start guessing "9111" through "9911", and so on and so forth.

An operation that's especially vulnerable to these so-called "timing attacks" is accessing memory. If systems always had to wait for memory before doing the next step of an action, they'd spend much of their time sitting idle.

To keep performance up, engineers employ a trick: they give the processor the power to execute multiple instructions while it waits for

memory—and then, once memory is ready, discards the ones that weren't needed. This is called "speculative execution."

While it pays off in performance speed, it also creates new security issues. Specifically, the attacker could make the processor speculatively execute some code to read a part of memory it shouldn't be able to. Even if the code fails, it could still leak data that the attacker can then access.

A common way to try to prevent such attacks is to split up memory so that it's not all stored in one area. Imagine an industrial kitchen shared by chefs who all want to keep their recipes secret. One approach would be to have the chefs set up their work on different sides—that's essentially what happens with the "Cache Allocation Technology" (CAT) that Intel started using in 2016. But such a system is still quite insecure, since one chef can get a pretty good idea of others' recipes by seeing which pots and pans they take from the common area.

In contrast, the MIT CSAIL team's approach is the equivalent of building walls to split the kitchen into separate spaces, and ensuring that everyone only knows their own ingredients and appliances. (This approach is a form of so-called "secure way partitioning"; the "chefs", in the case of cache memory, are referred to as "protection domains.")

As a playful counterpoint to Intel's CAT system, the researchers dubbed their method "DAWG", which stands for "Dynamically Allocated Way Guard." (The "dynamic" part means that DAWG can split the cache into multiple buckets whose size can vary over time.)

Lebedev co-wrote a new paper about the project with lead author Vladimir Kiriansky and MIT professors Saman Amarasinghe, Srini Devadas and Joel Emer. They will present their findings next week at the annual IEEE/ACM International Symposium on Microarchitecture (MICRO) in Fukuoka City, Japan.

"This paper dives into how to fully isolate one program's side-effects from percolating through to another program through the cache," says Mohit Tiwari, an assistant professor at the University of Texas at Austin who was not involved in the project. "This work secures a channel that's one of the most popular to use for attacks."

In tests, the team also found that the system was comparable with CAT on performance. They say that DAWG requires very minimal modifications to modern operating systems.

"We think this is an important step forward in giving computer architects, cloud providers and other IT professionals a better way to efficiently and dynamically allocate resources," says Kiriansky, a Ph.D. student at CSAIL. "It establishes clear boundaries for where sharing should and should not happen, so that programs with sensitive information can keep that data reasonably secure."

The team is quick to caution that DAWG can't yet defend against all speculative attacks. However, they have experimentally demonstrated that it is a foolproof solution to a broad range of non-speculative attacks against cryptographic software.

Lebedev says that the growing prevalence of these types of attacks demonstrates that, contrary to popular tech-CEO wisdom, more information-sharing isn't always a good thing.

"There's a tension between performance and security that's come to a head for a community of architecture designers that have always tried to share as much as possible in as many places as possible," he says. "On the other hand, if security was the only priority, we'd have separate computers for every program we want to run so that no information could ever leak, which obviously isn't practical. DAWG is part of a growing body of work trying to reconcile these two opposing forces."

It's worth recognizing that the sudden attention on timing attacks reflects the paradoxical fact that computer security has actually gotten a lot better in the last 20 years.

"A decade ago software wasn't written as well as it is today, which means that other attacks were a lot easier to perform," says Kiriansky. "As other aspects of security have become harder to carry out, these microarchitectural attacks have become more appealing, though they're still fortunately just a small piece in an arsenal of actions that an attacker would have to take to actually do damage."

The team is now working to improve DAWG so that it can stop all currently known speculative-execution attacks. In the meantime, they're hopeful that companies such as Intel will be interested in adopting their idea—or others like it—to minimize the chance of future data breaches.

"These kinds of attacks have become a lot easier thanks to these vulnerabilities," says Kiriansky. "With all the negative PR that's come up, companies like Intel have the incentives to get this right. The stars are aligned to make an approach like this happen."

More information: eprint.iacr.org/2018/418.pdf

Provided by Massachusetts Institute of Technology