Foreshadow in VMware-Umgebungen

Überschattet

Jörg Riether

Dass sich mit der Angriffsmethode Foreshadow ausgerechnet die Sicherheitstechnik SGX aushebeln lässt, ist vor allem bitter für Intel. Dass aber vor der NG-Variante VMs und Hypervisoren nicht mehr sicher sind, trifft vor allem die mit Virtualisierung arbeitenden Rechenzentren.

Als Anfang 2018 zwei Teams internationaler Sicherheitsforscher unabhängig voneinander die Schwachstelle Foreshadow entdeckten, die sich ähnlich wie Spectre und Meltdown der spekulativen Ausführung heutiger Intel-CPUs kombiniert mit einer Seitenkanalanalyse bedient, beschränkte sich die Reichweite auf Intels SGX-Technik. Durch die inzwischen unter CVE-2018-3615 gelistete Lücke erlangten die Teams Zugang zu in einer SGX-Enklave geschützten Daten und Kryptoschlüsseln im Klartext, indem sie sie in den Datenzwischenspeicher der CPU-Kerne L1d, kurz L1-Cache, kopieren und auslesen konnten. Die Links zur Dokumentation und zu allen weiteren Beschreibungen sind unter ix.de/ix1810098 zusammengefasst.

Intels Überprüfung brachte jedoch ans Licht, dass abgewandelte Techniken aka Foreshadow-NG nicht nur die von den Forschern erbeuteten Daten aus dem L1-Cache extrahieren können, sondern jede Information, die sich im L1-Cache des entsprechenden CPU-Kerns befindet. Mit dieser unter CVE-2018-3620 gelisteten Methode gelang es, auch den Speicher des Kernels auszulesen und den Speicher, den die im SMM (System Management Mode) ausgeführte Firmware verwendet.

Teilen sich virtuelle Maschinen mit anderen VMs oder mit dem Hypervisor einen physischen CPU-Kern, etwa weil Intels Hyperthreading ihn in zwei logische zerlegt, lassen sich deren Speicherbereich die Informationen mit der Variante CVE-2018-3646 entreißen. Intel selbst fasst alle drei unter L1TF (L1 Terminal Fault) zusammen.

Insbesondere für Betreiber und Kunden von Cloud-Infrastrukturen mit Virtualisierung ist die letztgenannte Variante von erheblicher Bedeutung. Foreshadow Typ 1 und 2 lassen sich mit Microcode-Updates, kombiniert mit Betriebssystem-Updates, Service Packs oder Kernel-Patches, abmildern. Auch dürften diese Updates inzwischen weit verbreitet sein. Bei Variante 3 ist es leider nicht so einfach. Zudem kann es erhebliche Konsequenzen für die Performance der Systeme haben.

Appliances betroffen

Im Folgenden sollen die Auswirkungen aller drei Varianten auf VMwares vSphere-Umgebungen beleuchtet werden. Die erste stellt keine Gefahr dar, da ESXi SGX weder benutzt noch für die Gäste virtualisiert. Anders bei Variante 2: VMware listet auf seiner Website alle virtuellen Appliances auf, für die bisher noch keine Patches vorhanden sind. Dazu zählen vCenter Server Appliance (vCSA) 6.0, 6.5 und 6.7.

Produkte, die nicht betroffen seien, listet VMware in seiner Knowledge Base auf. Die Begründung: Es gebe keinen Weg zum Ausführen von beliebigem Code ohne Administratorrechte, solange man sich an die empfohlene Konfiguration halte. Im selben Dokument heißt es, dass VMwares Hypervisoren ebenfalls nicht betroffen seien von CVE-2018-3620 alias Foreshadow Typ 2.

Die dritte Variante betrifft vSphere-Umgebungen allerdings erheblich. Sie erfordert Änderungen am Hypervisor-System. Die CVE-2018-3646 skizziert zwei Angriffsvektoren. VMware beschreibt diese als „Sequential-context Attack Vector“ und „Concurrent-context Attack Vector“. In beiden teilen sich zwei VMs respektive VM und Hypervisor einen Prozessorkern und damit seinen L1-Cache, in der Regel, weil ein eingeschaltetes Hyperthreading ihn als zwei logische Kerne durchreicht und der Hypervisor sein eigenes Scheduling der vCPUs darüberlegt.

Bei Ersterem kann ein Individuum oder System, das eine VM kontrolliert, an L1-Cache-Daten eines vorherigen Thread des Hypervisors oder einer anderen VM kommen, so sie auf dem gleichen Prozessorkern läuft. Dazu muss noch nicht einmal Hyperthreading aktiviert sein. Die Gefahr, dass jemand auf diese Weise an Daten kommt, kann man mit einem von Intel entwickelten Patch reduzieren, der ein „L1 Cache Flush“ betreibt. Er leert nach wie auch immer optimierten Regeln periodisch den L1-Cache und reduziert damit die Wahrscheinlichkeit eines erfolgreichen Angriffs. Den Patch stellte Intel von März bis Mai 2018 als CPU-Microcode-Update bereit.

Beim „Concurrent-context Attack Vector“ führen die Nutzer der beiden logischen CPUs dank Hyperthreading ihre Prozesse quasiparallel aus. Einen Zugriff des einen Thread auf die Daten des anderen will Intel deshalb mit dem „Core Scheduling“ unterbinden. Diese Technik erlaubt die quasiparallele Ausführung der beiden logischen CPUs nur unter bestimmten Umständen. Dadurch sind die Betriebssystemhersteller gezwungen, solche vertrauenswürdigen Rechenzeitnutzer – VMs oder Hypervisor – zu definieren, damit sie Zugriff auf denselben physischen Kern bekommen.

VMware stellt sie unter dem Namen „ESXi Side-Channel-Aware Scheduler“ oder kurz „SCA-Scheduler“ bereit. Die zum Redaktionsschluss vorliegende erste Version vollbringt aber wahrlich keine Wunder. Sie gewährleistet lediglich, dass immer nur ein logischer Prozessor eines Hyperthreading-aktivierten Kerns benutzt werden kann, gleich ob von einer VM oder vom Hypervisor selbst. Vertrauenswürdige CPU-Nutzer sind noch nicht implementiert. Das wirft die Frage auf, warum man Hyperthreading nicht gleich abschaltet. VMware rät davon ab, weil es ja sein könne, dass man es dank zukünftiger Verbesserungen wieder aktivieren müsse. Die Leistungseinbußen erörtert VMware in seiner Knowledge Base.

VMware liefert den Patch über den Update-Manager aus, aktiviert aber nur den Patch mit dem L1 Cache Flush, sofern das entsprechende Microcode-Update vorhanden ist. Die Leistungsverluste beziffert VMware mit maximal 3 %. Den SCA-Scheduler installiert VMware zwar, schaltet ihn aber nicht scharf. Der Grund dürfte in den erheblichen Leistungseinbußen liegen, die eintreten können. Im ungünstigen Fall, bei einem voll ausgelasteten Host mit wenig oder gar keinen CPU-Reserven, kann die Systemperformance nach der Aktivierung des SCA-Scheduler laut VMware um 32 % zurückgehen (siehe Tabelle „Performanceverluste durch den SCA-Scheduler“).

In voll ausgelasteten Umgebungen kann es nach der Aktivierung des SCA-Scheduler zu erheblichen Leistungseinbußen kommen.

Mögliche Performanceverluste durch den SCA-Scheduler laut VMware-Labor

Anwendungs-Workload/Gast-OS	Performanceeinbußen nach Aktivieren des SCA-Scheduler
Datenbank OLTP/Windows	32 %
Datenbank OLTP/Linux (mit vSAN)	32 %
gemischte Workloads/Linux	25 %
Java/Linux	22 %
VDI/Windows	30 %

Mit weniger Leistung

In vSphere-Umgebungen mit großzügiger CPU-Reserve sollte man auch mit aktiviertem SCA-Scheduler über die Runden kommen. VMware rechnet vor, dass bei einem Linux-OLTP-Datenbanksystem und einem Hypervisor-Host, dessen CPU-Gesamtlast bei etwa 62 % liege, der Leistungsverlust nach Aktivieren des SCA-Scheduler nur etwa 1 % betrage. Aber Obacht: War die identische Maschine vor der Aktivierung zu 90 % ausgelastet, liege der Leistungsverlust danach bei 32 %.

Auch wer im eigenen Rechenzentrum über ausreichend Reserven verfügt, sollte sich die Lastentwicklung über einen längeren Zeitraum vor Augen führen. Es genügt, wenn in einer großzügig ausgelegten Umgebung mit Hunderten ESXi-Hosts zu einem Zeitpunkt bestimmte Tasks hundert- oder tausendfach gleichzeitig laufen, seien es nun zahlreiche Datenbanktransaktionen, Backup-Aufträge oder Windows-Updates. In solchen Momenten könnte selbst eine zuvor wie geschmiert laufende Infrastruktur plötzlich haken.

Man sollte also seine Umgebung und insbesondere die CPU-Last ausführlich analysieren, bevor man den SCA-Scheduler einschaltet. Dies kann man in den erweiterten Systemeinstellungen, indem man die Variable *VMkernel.Boot.hyperthreadingMitigation* auf *true* setzt und anschließend einen Reboot durchführt.

VMware hat zudem ein Werkzeug entwickelt, mit dessen Hilfe man eine solche Evaluierung automatisieren kann (siehe ix.de/ix1810098). Schaltet man den Scheduler nicht scharf, gibt vSphere für die betreffenden Hosts eine Warnmeldung aus, die sich auf Wunsch deaktivieren lässt, indem man die Variable *UserVars.SuppressHyperthreadWarning* auf *1* setzt. Obgleich Intel darauf hinweist, dass es sich bei L1TF um eine hochentwickelte Angriffsmethode handele und bis heute keine Berichte über tatsächliche Angriffe bekannt seien, sollte man dies dennoch ernst nehmen und seine Umgebung nach den jeweiligen Möglichkeiten absichern, vor allem, wenn Externe Zugriff darauf haben.

In der VMware Cloud on AWS ist laut VMware der sequenzielle Angriffsvektor bereits „mitigated“, also abgeschwächt, die Anpassung für den gleichzeitigen Angriffsvektor gelang kurz vor Redaktionsschluss. Für die Horizon Cloud werden entsprechende Updates derzeit priorisiert entwickelt. VMware Workspace One SaaS sei auf Grundlagen der derzeitigen Bewertungen von Foreshadow nicht betroffen.

Die Workstation und den Player hat VMware ab Version 14.1.3, die Mac-Hypervisor Fusion und Fusion Pro ab 10.1.3 mit Updates zum verbesserten Schutz gegen den sequenziellen Angriffsvektor ausgestattet. In Hinblick auf den gleichzeitigen Angriffsvektor empfiehlt der Hersteller, Hyperthreading komplett zu deaktivieren. An einem PC geht man dafür ins BIOS- oder EFI-Setup, für den Mac hat VMware ein Werkzeug bereitgestellt (siehe ix.de/ix1810098).

Fazit

Zwar kann man seine virtuellen Umgebungen mit Intel-CPUs gegen Foreshadow-Angriffe schützen, doch muss man sich vor allem bei solchen mit einer sehr hohen CPU-Gesamtauslastung auf Leistungseinbußen einstellen. Gleichzeitig bleibt zu hoffen, dass es VMware gelingt, den SCA-Scheduler performanter zu gestalten. (sun@ix.de)

Jörg Riether

ist spezialisiert auf die Bereiche IT-Sicherheit, Hochverfügbarkeit und Virtualisierung. Er arbeitet als Leiter der IT bei der Vitos Haina gemeinnützige GmbH.

Alle Links: ix.de/ix1810098