Issued 5/10/18

Every Major OS Maker Misread Intel's docs. Now Their Kernels can be Hijacked or Crashed

Linux, Windows, macOS, FreeBSD, and some implementations of Xen have a design flaw that could allow attackers to, at best, crash Intel and AMD-powered computers. At worst, miscreants can, potentially, "gain access to sensitive memory information or control low-level operating system functions," which is a fancy way of saying peek at kernel memory, or hijack the critical code running the machine. The vulnerabilities can be exploited by malware running on a computer, or a malicious logged-in user. Patches are now available to correct the near-industry-wide programming blunders. As detailed by CERT on Tuesday, the security issue, labeled CVE-2018-8897, appears to have been caused by developers at Microsoft, Apple, and other organizations misunderstanding the way Intel and AMD processors handle one particular special exception. Indeed, CERT noted: "The error appears to be due to developer interpretation of existing documentation." In other words, programmers misunderstood Intel and AMD's manuals, which may not have been very clear. At the heart of the issue is the POP SS instruction, which takes from the running program's stack a value used to select the stack's segment, and puts that number into the CPU's stack selector register. This is all to do with memory segmentation that modern operating systems mostly ignore, and you can, too. The POP SS instruction is specially handled by the CPU so that the stack cannot be left in an inconsistent state if an interrupt fires while it is executing. An application can set a debug breakpoint for the memory location where that stack selector will be pulled from the stack by

POP SS. That is, when the app uses POP SS, it will generate a special exception when the processor touches a particular part of RAM to fetch the stack selector. Now, here's the clever trick. The instruction immediately after the POP SS instruction has to be an INT instruction, which triggers an interrupt. These software-generated interrupts are sometimes used by user programs to activate the kernel so it can do work for the running process, such as open a file. On Intel and AMD machines, the software-generated interrupt instruction immediately after POP SS causes the processor to enter the kernel's interrupt handler. Then the debug exception fires, because POP SS caused the exception to be deferred. Operating system designers didn't expect this. They read Intel's x86-64 manuals, and concluded the handler starts in an uninterruptable state. But now there's an unexpected debug exception to deal with while very early inside the interrupt handler. Red Hat has patches ready to roll, as does Ubuntu, and Apple for macOS. The Linux kernel has also been fixed, way back on March 23, 2018. A patch is already present in versions 4.15.14, 4.14.31, 4.9.91, 4.4.125, plus older 4.1, 3.16, and 3.2 branches. Microsoft’s got it sorted, for Windows 7 through 10 and Windows Server 2008 through version 1803. Xen has patches for versions 4.6 through 4.10. VMware’s hypervisors aren’t at risk, but vCenter Server has a workaround and vSphere Integrated containers await a fix, but both are rated merely “potentially affected.”