IDC PERSPECTIVE

Intel Provides New Tools to the Cybersecurity Task: The Results Could Be Game Changing

Frank Dickson Michael Suby

EXECUTIVE SNAPSHOT

FIGURE 1

Executive Snapshot: Intel Threat Detection Technology

Under the moniker Intel Threat Detection Technology (TDT), Intel is unlocking capabilities in its system on a chip offerings that fundamentally change the ability of security vendors to implement security. Intel TDT leverages silicon-level telemetry and functionality that security independent software vendors and possibly original equipment manufacturers can leverage to enable the hardware compute platform to play an active role in threat defense against "above the OS" attacks such as ransomware and cryptomining.

Key Takeaways

- The goal of the Intel/security software vendor partnership is to enable and empower the Intel-based systems of today and tomorrow to be fundamentally more secure and have lower malware infection rates than AMD, Apple, and other ARM-based processor systems.
- Intel TDT has different capabilities, and most are backward compatible to sixth-generation Skylake processors. The Tiger Lake generation adds new multidetector capabilities and performance increases.
- Adoption could lead to a measurable difference in security efficacy between Intel-based systems and the systems based on other processor vendors.

Recommended Actions

- The selection and procurement of personal computing devices and server systems have a long-term, measurable impact on the security posture of an organization, as they are not commodities. Procurement, IT, and security should begin to have discussions on how devices will affect the hard and soft costs of security.
- Requests for proposal (RFPs) should include specifications that consider the impact of security.
- Software vendors should discuss how to leverage new features to improve the security of their offerings.

Source: IDC, 2021

March 2021, IDC #US47495021

SITUATION OVERVIEW

Security as a differentiator in hardware-based computing platforms has largely been limited to attributes that live "below the operating system (OS)," as the root-of-trust capabilities are typically enabled by computing device platform providers and operating system vendors to improve system integrity. A best practice for IT and security teams is to start with PCs and servers built on the foundation of hardware-based security. Burned into silicon, root of trust checks the authenticity of the firmware and OS at the boot and during the operation. If alterations are detected, golden images are automatically returned. Cybersecurity is improved, and the occurrences of disruptive out-of-commission PCs are reduced. It's important to note that security features in silicon can enable additional security features such as trusted enclaves, secure key storage, crypto acceleration, and isolation for apps and virtual machines.

The world above the OS is largely unaffected by chip vendors. This world is governed by software security measures hosted on top of the OS. From chips up to devices, security differentiation among device manufacturers (i.e., original equipment manufacturers [OEMs]) below, at, and above the OS is limited but has been expanding. Historically, Apple Macs have been the productivity platform of choice by some security professionals as these devices have been viewed as more secure. That perception, though, may be more due to the fact that Macs are less common than Windows-based systems, so less malware is written to target them. HP Inc. and Google are examples of device manufacturers with a holistic vision on device security. HP Business PCs include security capabilities (e.g., Sure Start, Sure Run, Sure Recover, Sure Click, and Sure Sense) that perform security functions below, at, and above the OS. Google Chromebooks include hardware-assisted security features that also span below, at, and above the OS. Finally, Microsoft's Windows 10 secured-core PCs, built in partnership with Microsoft OEM partners, is another example of devices with security built on top of silicon-enabled security capabilities as a foundation.

At the chip level, under the moniker of Intel TDT, Intel is unlocking capabilities in its system on a chip (SoC) that fundamentally change the rules of the game. Intel TDT is a capability that is part of the Intel Hardware Shield advanced threat detection category. The other recently released new capability in this category is Intel Control-Flow Enforcement Technology (CET), which closes the door on memory-related control flow hijack attacks.

Intel's strategy leverages silicon-level telemetry and functionality that security independent software vendors (ISVs) and possibly OEMs can leverage to enable the hardware compute platform to play an active role in threat defense against "above the OS" attacks such as ransomware and cryptomining. The goal of the Intel/security software vendor partnership is to enable and empower the Intel-based systems of today and tomorrow to be fundamentally more secure and have lower malware infection rates than AMD, Apple, and other ARM-based processor systems.

Intel TDT has different capabilities, and most are backward compatible to sixth-generation Skylake processors. The Tiger Lake generation adds new multidetector capabilities and performance increases.

Applying the PMU to Security Use Cases

The performance monitoring unit (PMU) enables developers to monitor and tune app performance inside Intel processors, providing a precise and dynamic picture for how granular application processes are performing at the central processing unit (CPU) level. The performance and compute capacity of today's processors can be heavily impacted by the manner in which applications leverage hierarchical cache subsystems, nonuniform memory, simultaneous multithreading, and out-of-order

execution. Applications with granular performance insights can be adjusted for resource utilization to improve performance and power usage.

Why Does the PMU Matter to Security?

The recent innovations in malware tend to be very compute intensive. Cryptojacking, for example, installs malware on a device or server to steal computing resources for the mining of cryptocurrency (referred to as cryptomining). Cryptomining essentially performs the same compute-intensive encryption repeatedly.

Ransomware is another example of malware. Although ransomware has evolved to do many malicious things, at its core, ransomware performs the same encryption operations typically based on three main encryption algorithms in quick succession in accessing, encrypting, and saving a file and deleting the backup. Such operations are consistent (nonvariable) and compute intensive. By using the PMU to observe the exact manner in which any program executes at the CPU level, security ISVs can illuminate stealthy malware.

The idea behind this technology is to profile how malware executes on the PMU. The concept is to first extract a processes-level execution fingerprint and then use that for detecting malware by using machine learning (ML) to compare the fingerprint to a known library of malware. Intel helps develop the right signatures that security software vendors integrate into endpoint agents as additional high-fidelity signals to analyze with their other file- or behavior-based methods to identify the threats.

BlackBerry is the first security vendor, to IDC's knowledge, that has leveraged this new PMU feature, addressing the sophisticated cryptojacking malware detection use case. BlackBerry's Optics Context Analysis Engine, which it acquired as part of the Cylance acquisition, utilizes PMU telemetry to identify elusive cryptojacking attempts. Illusive cryptojacking approaches may distribute small quantities of compute across a multitude of machines to create a pool of processing power in the hopes of avoiding detection. The BlackBerry implementation for cryptomining detection is not only effective but efficient as PMU telemetry-based detection has very little processor impact and is straightforward to configure. Let's face it — IT administrators concerned with PC performance are hypersensitive to the compute overhead of new security features, especially as it relates to cryptomining detection as some detection approaches can be math (CPU compute) intensive.

The PMU detection approach is not only about detecting malware but also about *not* falsely flagging applications as malware. Common tasks performed by a PC or server can resemble malware behavior without the appropriate context. For example, opening a Word document, editing the text, adding a graphic, saving the file, and closing the document are common activities. Without context, a cybersecurity solution may just see the accessing of a file, the dramatic increase of the file size, the saving of the file, and the deletion of the backup copy. That context-free view resembles ransomware. By incorporating CPU execution behavioral analysis and other context from the PMU, the cybersecurity solution would see that the application did not have the match of how a particular ransomware encryption process would touch the same PMU. Thus the cybersecurity solution would have the superior context to eliminate a false positive.

The approach to discovering ransomware is not necessarily new. EDR vendors do signature-based, file-based scanning and file behavioral analysis. CPU telemetry based on the PMU adds an entirely new class of behavior monitoring based on the CPU process itself. The approach provides a net-new signal that was not previously available for consideration. Intel has modeled and tested against 50

common ransomware strains. The breadth of research also enables the detection of zero-day variants, since they derive from the same encryption classes.

Giving Security Vendors the Keys to the GPU

PMU for detection is super cool and adds real value, especially given its resource efficiency. However, it does not have the same potential impact of leveraging graphics processing units (GPUs) in Intel processors. Applying GPUs to security use cases can be game changing, in IDC's opinion.

Malware detection is essentially a math-intensive science. As mainstream security ISVs have look to empower their detection engines with artificial intelligence (AI) and machine learning, security ISVs say "welcome to the party," as they have been leveraging AI and ML technologies for decades. As malware detection moved from signature-based pattern matching to real-time static and behavioral algorithmic-based detection, the science became more compute intensive, demanding more compute from end devices.

This need for compute creates a problem. First, a general-purpose CPU is good for general-purpose, logic-based tasks. However, CPUs are not good at math-heavy, repeatable (parallel)-based tasks. Other processor types such as a digital signal processor or a GPU are better at those. Think of it like this: A pickup truck is a very good general-purpose vehicle that can do many things. You can make a pickup truck go fast, but if you want to go really fast with great acceleration, a sports car would be much better suited.

Second, although security and malware detection use cases can have an insatiable appetite for compute, datacenter and IT administrators and end users do not necessarily want to "feed the beast." Complex, compute-heavy application requirements imply the need of additional compute for datacenter managers. For end-user devices, compute-intensive applications create poor user experiences as security slows system performance and users have to wait.

The insatiable appetite for compute has resulted in artificially limiting the compute tax of the ISV's security applications to prevent poor user experiences. For endpoint protection platforms, you will commonly hear vendors refer to their "lightweight" agent. The informal and unspoken limit to CPU utilization by endpoint protection platforms (EPPs) seems to be set at 2%, achieved by constraining the complexity and sophistication of detection algorithms. Many security vendors look to circumvent the endpoint utilization caps by doing much of the compute heavy lifting in the cloud. CrowdStrike is normally acknowledged as one of the early innovators of this approach that has now been adopted by other EPP ISVs.

Why Does the GPU Matter to Security?

Opening the GPU to security workloads provides an exponential increase in math-centric compute that can be applied to security use cases. Intel TDT provides an API for security software vendors to offload any security workload to the GPU. Even Intel's Accelerated Memory Scanning (AMS), which was released by Intel TDT in 2018, offloaded to the GPU for this capability. New releases added a number of new AI/ML-based, pre-canned offload types that can be offloaded to the GPU, and Intel made this customizable to be leveraged in different ways. This new resource enables security vendors to deliver three potential benefits.

First, compute can be offloaded from the CPU to the GPU. The 2% cap can be used for other tasks. Remember, the GPU is not only better at compute for many security workloads but also ripe for being leveraged, especially in corporate use cases. A GPU in a commercial scenario is relatively unused, as

opposed to a consumer setting where gaming and other apps highly leverage the GPU. Therefore, there is a nice match between unused horsepower and the need for security in commercial.

Second, and significantly more important, the "handcuffs" can be removed from security providers as to the complexity of detection algorithms used. Models that were too complex and violated the upper limit of the 2% cap can now be leveraged. The improved efficacy not only detects malware that may have gone undetected in the past but also reduces the number of false positives.

The first vendor, to IDC's knowledge, that is leveraging the GPU ransomware detection is Cybereason. Technically, Blackberry's cryptomining detection also leverages Intel TDT's offload to the GPU. Cybereason is first to announce for ransomware.

Cybereason is leveraging the GPU capabilities unleashed under the moniker Intel Hardware Shield protections to address ransomware detection. Cybereason can now leverage GPUs to accelerate EPP, performance-intensive ML security algorithms by offloading tasks from the CPU to the Intel integrated graphics controller.

The Cybereason use case also uses the PMU telemetry detailed previously. As opposed to the BlackBerry implementation that comes with an imperceptible resource tax for cryptomining, the processing power to use machine learning to scan for different strains of ransomware is a high-compute process. Thus, for ransomware detection, Cybereason offloads the machine learning models from the CPU to the GPU.

Finally, leveraging the power of the GPU can unleash completely new use cases that would not have been possible in the past. Memory scanning can be an extremely effective way to discover elusive malware, but memory scanning can also be resource intensive and slow a device to a "crawl." The impact to user experience may force a compromise to limit active memory scanning to once an hour, once a day, or maybe not at all. By offloading memory scanning to the GPU, memory scanning can be done almost continually without any perceptible impact to system performance as perceived by the user. Intel's Accelerated Memory Scanning feature, initially demonstrated on benchmark testing, dropped CPU utilization from 20% to as little as 2% for memory scanning use cases. One might say that it provides another arrow in the quiver of the security vendor; it might be more accurate to describe it as upgrading from a bow and arrow to a rifle.

The first vendor, to IDC's knowledge, to leverage the memory scanning capability is Microsoft. Microsoft shipped Accelerated Memory Scanning (GPU offload for pattern matching) in 2018 as a part of Microsoft Defender for Endpoint. Sentinel One later announced the integration of Accelerated Memory Scanning capabilities with the Sentinel One autonomous endpoint protection platform on February 27, 2019.

ADVICE FOR THE TECHNOLOGY BUYER

Clearly, IDC is fan of the strategy. The features provided by Intel enable a demonstrable improvement to security.

The key to the strategy is security vendor utilization. So far, adoption appears to be light. Announcements from BlackBerry, Cybereason, Microsoft, and Sentinel One are promising. Given that Intel Threat Detection Technology was announced on April 16, 2018, we would have expected more.

In fairness, leveraging the features requires new data research and code enablement. Still, the key to the strategy is broad security vendor adoption. "Light" would be the appropriate adoption description.

Also, Intel intentionally targeted specific vendors to pilot features such as Accelerated Memory Scanning to optimize the detectors production quality tuned for wider implementation. These technologies take time to validate the increased efficacy for the security software vendors and then integrate them into these complex security stacks.

A concern has been communicated to IDC about leveraging Intel TDT as a variance in the level of protection provided could vary based on the processor type of the device. Consistent levels of protection, regardless of the device and processor type, is the goal of security vendors. Although IDC acknowledges the concern, it would seem to validate Intel's strategy. Superior and differentiated cybersecurity protection of Intel-based devices is Intel's goal after all.

In spite of the tepid current adoption, the potential is huge. Accelerated adoption could lead to a measurable difference in security efficacy between Intel-based systems and the systems based on other processor vendors. The implication is that selection and procurement of personal computing devices and server systems will have a long-term impact on the security posture of an organization.

The operative word in the previous paragraph is potential. The vision is fantastic, but little current data exists to confirm or deny whether the vision will be fulfilled.

LEARN MORE

Related Research

- *Insights from IDC's EDR and XDR 2020 Survey: Operational Challenges and Initiatives Are Abundant* (IDC #US47357921, January 2021)
- *Organizations Are Crying Out for Assistance in Measuring IT and Security* (IDC #US46984720, November 2020)
- *Chromebook's Impact on the Endpoint Security Market Not to Be Dismissed* (IDC #US45637419, November 2019)

Synopsis

This IDC Perspective discusses how Intel, under the moniker Intel Threat Detection Technology (TDT), is unlocking capabilities in its system on a chip (SoC) that fundamentally change the ability of security vendors to implement security. Intel TDT leverages silicon-level telemetry and functionality that security independent software vendors (ISVs) and possibly OEMs can leverage to enable the hardware compute platform to play an active role in threat defense against "above the OS" attacks such as ransomware and cryptomining.

"Unlocking silicon-based features, the GPU for security workloads is a no-brainer. The GPU is not only better at compute for many security workloads but also ripe for being leveraged, especially in corporate use cases. A GPU in a commercial scenario is relatively unused, as opposed to consumer where gaming and other apps highly leverage the GPU. Therefore, there is a nice match between unused horsepower and the need for security in commercial," according to Frank Dickson, program vice president, Security and Trust.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com