HARDWARE SECURITY AND CONFIDENTIAL COMPUTING IN SERVER PLATFORMS

Author:

Ralf Helkenberg

April 2023

An IDC Industry Spotlight sponsored by AMD

IDC #EUR150569623

Hardware Security and Confidential Computing in Server Platforms

Introduction

IDC research shows that organizations are shifting to a digital-first strategy, where the focus is moving from iterative transformation to running a sustainable digital business. The business imperative to launch new digital services and applications quickly and easily, automate processes, and store and process increasing amounts of data is driving modernization of IT infrastructure. This includes supporting hybrid and multicloud architectures, AI, and Big Data analytics platforms to distributed data infrastructure models.

Delivering on the digital-first promise though has potentially expanded the attack surface that organizations need to monitor and protect from cyber risks. This is against a backdrop of an ever-changing cyberthreat landscape. Ransomware attacks have become increasingly sophisticated and widespread, causing major disruption and data loss for organizations of all sizes. The rapid move of business-critical workloads to the cloud has seen a rise in cloud-based attacks, and supply chains are emerging as a new vector of attack as attackers seek to exploit vulnerabilities in a third-party supplier to rapidly compromise an extended network of customers and partners. The growing business cost and impact to an organization's reputation from cyberattacks and data breaches is leading organizations to raise the security posture of their IT infrastructure so that they are better able to deliver the intended digital-first outcomes.

AT A GLANCE

KEY STATS

IDC European infrastructure and security survey findings:

» 55% of organizations saw an increase in the volume of cyberattacks over the prior 12 months.

» For 58% of organizations, building cyber resilience into infrastructure and processes is a strategic priority.

» Improving security and compliance posture is the number 1 datacenter priority for 2022–2023.

KEY TAKEAWAYS

» Security software is not enough protection from emerging threats. State-of-the-art hardware-based protection at the processor level should be considered to help increase overall system security posture.

» Confidential computing is unlocking new data-in-use compute scenarios, primarily to protect confidential VM workloads in the cloud.

The security of applications and data hosted in a datacenter or in the cloud depends partly on how secure the hosting server, storage, and networking infrastructure are. With the shift toward software-defined infrastructure, much of the focus and investment in security has been directed toward the application and networking layers. This potentially leaves business-critical data in use and data at rest exposed to hardware vulnerabilities. The growing incidence of low-level, hardware-centric attacks have underscored the need to treat infrastructure platform security with the same level of importance as software and data management security.

Infrastructure platform security begins at the microprocessor level. A design and default approach to security by processor manufacturers has led to processors being introduced into the market with security features that are better equipped to help protect against a range of threats and provide an added layer of infrastructure protection.

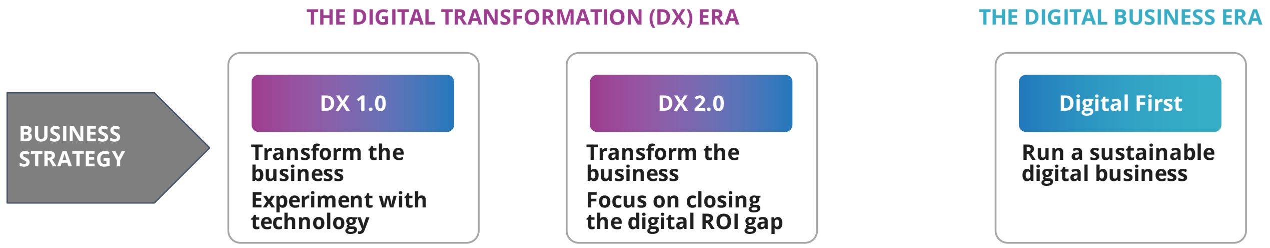
With advanced built-in security capabilities such as confidential computing, encrypted virtual machines (VMs), and containers, organizations have options to put in place hardware-based security to support their digital-first ambitions.

Computing Security Is Central to IT Infrastructure Transformation

1. From Digital Transformation to Digital First: Accelerating Digital Value

IDC research shows that 70% of organizations in Europe are looking for ongoing viable growth built on a digital-first strategy. The goal is to create new sources of value through digital products, digital services, and digital experiences. IDC is also observing a shift in organizational strategies and resources with respect to digital where the focus is moving from transformation to running a viable digital business (see Figure 1).

FIGURE 1
The Shift to the Digital Business Era

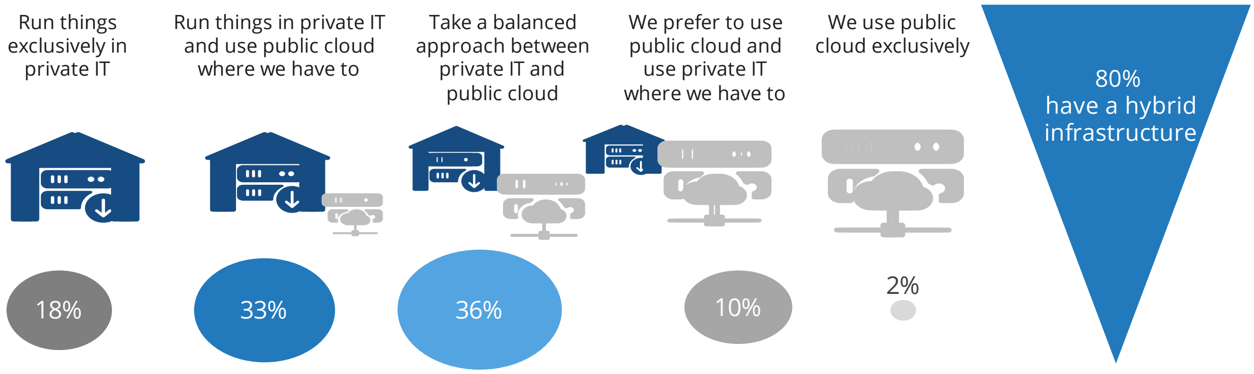

THE DIGITAL TRANSFORMATION (DX) ERA
THE DIGITAL BUSINESS ERA
BUSINESS STRATEGY
DX 1.0
Transform the business
Experiment with technology
DX 2.0
Transform the business
Focus on closing the digital ROI gap
Digital First
Run a sustainable digital business

Source: IDC Worldwide Digital Transformation Spending Guide, 2021 V2

To thrive in this emerging digital-first world, organizations are using cloud platforms to help drive innovation, modernization, productivity, and efficiency. IDC research shows that most of the European market relies on a hybrid approach to IT infrastructure and strongly prefers to deploy applications or services to private IT infrastructure rather than public cloud platforms (see Figure 2). As the figure shows, there are very few organizations that operate in a public-cloud-only business model. Rather, public cloud is a way to augment private IT infrastructure rather than replace it.

FIGURE 2

80% of European Organizations Have a Hybrid Approach to IT Infrastructure



Source: IDC's European Multicloud Survey, 2022

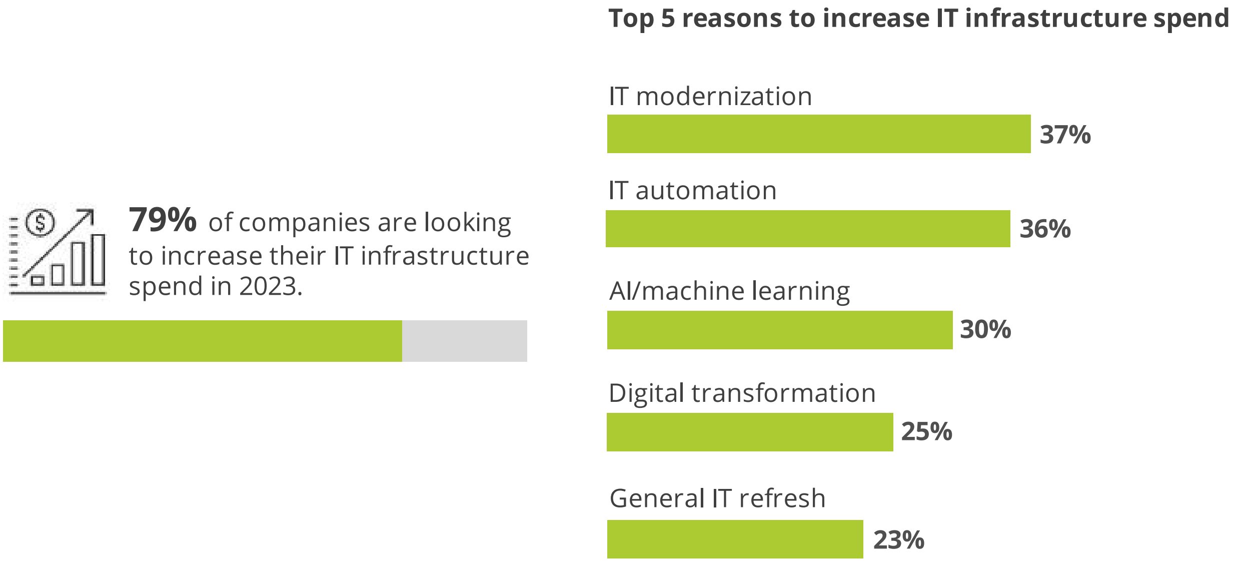
2. Shift to a Digital-First Strategy Is Driving Strong IT Infrastructure Investment

To support the shift to a digital-first strategy, European organizations are focused on investing in their IT infrastructure as data and digital services become a dominant source of business revenue and success. IDC research shows that over three-quarters of organizations are planning to increase spending on IT infrastructure in 2023 despite the uncertain economic climate (see Figure 3). The top 3 drivers for increased investment are:

- **IT modernization:** replacing legacy systems with a modern IT infrastructure that is more flexible, scalable, and adaptable to changing business needs
- **IT automation:** streamlining IT operations to improve overall infrastructure efficiency and resource utilization
- **AI/machine learning:** supporting artificial intelligence and machine learning applications, which require large amounts of data to train algorithms and models

FIGURE 3

Three-Quarters of Organizations Plan to Increase Spending on IT Infrastructure in 2023



Source: IDC's European Infrastructure Survey, 2022

3. Cyber Resilience: A Business Risk Mitigation Priority

Organizations anticipate that a wide range of external risks could affect their digital transformation and technology investment plans. According to IDC research, the main business risks cited by European business leaders are exposure to cyberattacks and non-compliance with data regulations (see Figure 4).

FIGURE 4

CEO Sentiment: Top 3 Risks with Greatest Impact on Business Today and in Two Years

2022	2024
1. Cybersecurity threats and regulations	1. Cybersecurity threats and regulations
2. Ensuring health/safety of employees and customers	2. Addressing new data sharing and compliance regimes
3. Operations resilience (e.g., supply chain risk)	3. Operations resilience (e.g., supply chain risk)

Source: IDC European Security Survey, 2022

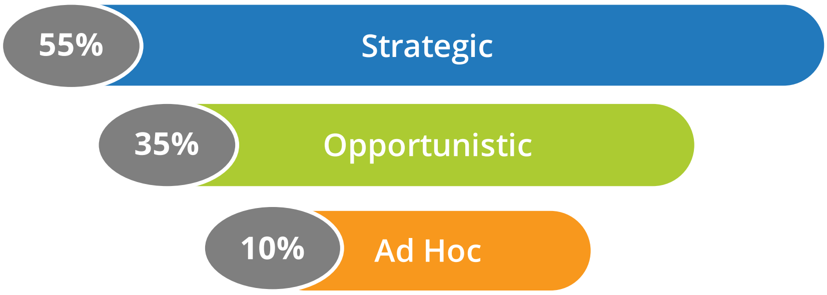
The cyberthreat landscape is constantly evolving, with new and more sophisticated threats emerging all the time. In IDC's *European Security Survey 2022*, 55% of respondents saw an increase in the volume of cyberattacks over the previous 12 months, with email compromise and ransomware the most cited incidents. The primary access vectors that threat actors use to get into an organization's environment are social engineering, brute-force credential attacks, and the exploitation of known software vulnerabilities. The business cost and impact from such cyber disruption are leading organizations to raise their security posture and strengthen their cyber resilience.

Cyber resilience goes beyond the more traditional view of business continuity or disaster recovery plans; it means understanding at a much broader level how cyberincidents or failures can impact the business across the entire value chain and planning how to avoid, reduce, or compensate for those failures. Given the growing dependency on cloud-based services and managed service providers, policy makers and industry are putting greater focus on cyber resilience of business IT infrastructure, particularly in national critical infrastructure sectors.

Cyberattacks on the update mechanisms of managed service providers Kaseya and Solarwinds highlighted how seemingly small players in the IT supply chain can introduce disproportionately high levels of cybersecurity risk to the customer environment. Industry and governments are therefore putting increased focus on the need for trustworthy supply chains, trustworthy partners, and trusted systems. IDC's *European Security Survey 2022* revealed that while cyber resilience is a strategic priority for most European organizations, the implementation of technologies and processes to help build such resilience still has some way to go (see Figure 5).

FIGURE 5

Cyber Resilience Is a Strategic Priority



Source: IDC European Security Survey, 2022

The ability to demonstrate satisfactory cyber resilience against a broader digital risk environment is also the focus of forthcoming European regulations:

- NIS2 (the Network and Information Security Directive) aims to improve the security and resilience of network and information systems across critical infrastructure in the EU. NIS2 significantly expands the sectors and type of entities falling under its scope, and organizations must take appropriate technical and organizational measures to manage cybersecurity risks and prevent and minimize the impact of potential incidents, including the effective use of encryption. Starting in January 2023, member states will have 21 months to transpose NIS2 into national law.
- The EU is also strengthening the digital resilience of the banking and financial services industry with the Digital Operational Resilience Act (DORA). All EU financial entities are required to ensure they can withstand, respond to, and recover from all types of ICT-related disruptions and threats. Though coming into force in January 2023, the new rules will apply from January 2025. DORA will be directly effective from January 17, 2025.

4. Secure Data Computation in the Public Cloud

Public cloud platforms are typically multitenant environments sharing the same infrastructure. This presents data security and privacy compliance concerns and is a key inhibitor to European organizations moving sensitive data into the public cloud.

Security concerns center around the perceived increased risk of data breaches and other security incidents, the lack of visibility into what data is held within cloud applications, and the extent to which organizations have control over who can access sensitive data, especially cloud service provider employees or contractors — and latterly the extent to which government and law enforcement can access and request customer data.

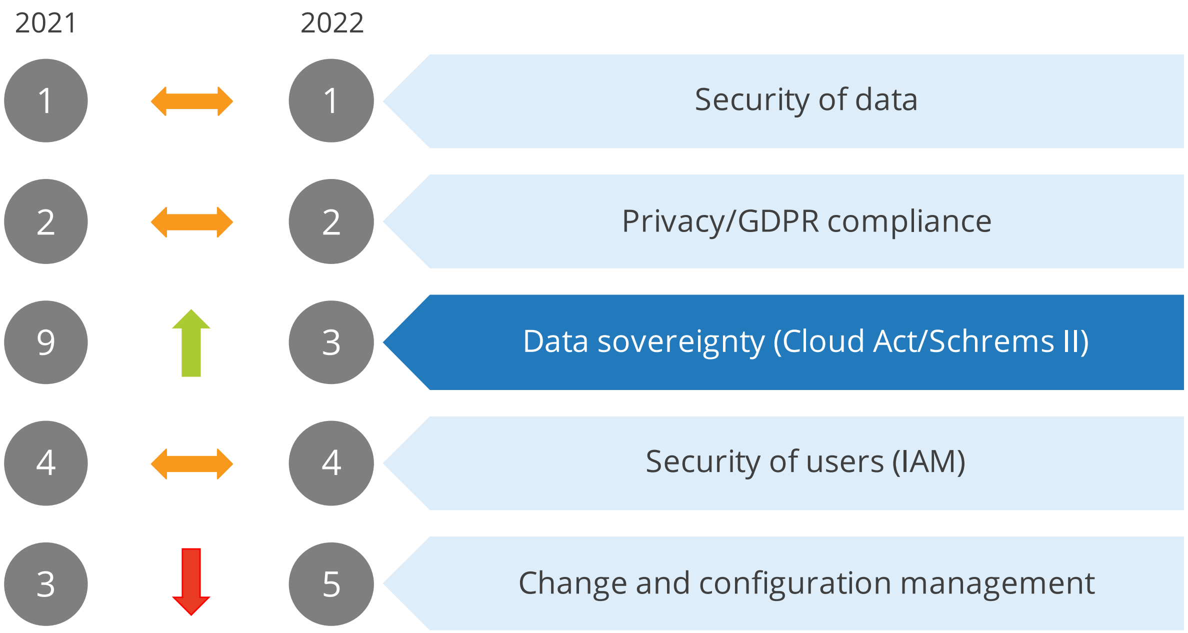
More data protection regulations, either modeled closely on the General Data Protection Regulation (GDPR), or with strong standards for protection, are emerging across the globe. Countries are taking more sovereignty measures to control the infrastructure and data generated in their jurisdictions, from rules on data residency to conditions on transborder data flows.

The growing extraterritorial application of data governance laws subjects organizations to growing tension between allowing digital innovation to accelerate and ensuring data and IT infrastructures comply with regulations and guidelines. The data sovereignty implications extend to the cloud environment given that organizations are increasingly moving their services and data to platforms managed by international cloud providers.

IDC research shows data sovereignty becoming a leading trust and compliance concern for European organizations when shaping their cloud strategy. Ranked only ninth in IDC's *2021 European Security Survey*, data sovereignty rose to third in terms of importance in 2022, behind privacy compliance and data security (see Figure 6).

FIGURE 6

Trust Concerns Impacting Cloud Strategy



2021		2022	
1	↔	1	Security of data
2	↔	2	Privacy/GDPR compliance
9	↑	3	Data sovereignty (Cloud Act/Schrems II)
4	↔	4	Security of users (IAM)
3	↓	5	Change and configuration management

Source: IDC European Security Survey, 2022

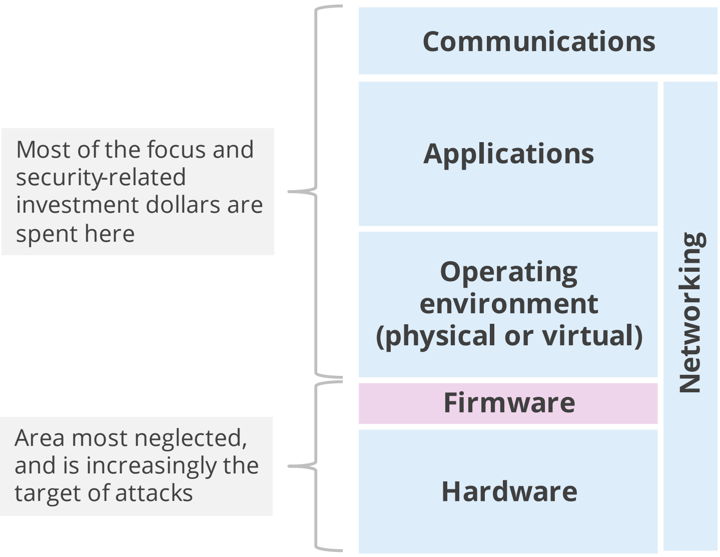
Trusted Computing: Hardware-Based Security and Confidential Computing

Protecting Server Platforms from Cyberthreats

Protecting users, data, and systems requires strong alignment between hardware and software technologies. However, given the increasing complexity of software and the rapidly evolving threat landscape, software security has tended to be a higher-priority concern for organizations.

However, hardware security should not be overlooked. Cyberattacks on hardware and firmware layers are becoming more prevalent as attackers seek new ways to access important information and credentials, gain remote control of systems, and cause disruption — a favorite vector for ransomware attacks, where hackers lock up a system remotely through root access.

Occurring below the operating system, such attacks can be difficult to detect because traditional security measures to detect and quarantine malware, such as antivirus software, may not detect them.



Most of the focus and security-related investment dollars are spent here

Area most neglected, and is increasingly the target of attacks

These attacks can take many forms, including:

- **Secure boot attacks:** Introduce malicious code into the system by bypassing or compromising the secure boot process.
- **Firmware attacks:** Target vulnerabilities in the hardware's firmware, enabling attackers to gain control of the system at a low level.
- **Dynamic random access memory (DRAM) threats:** Malware exploitation of DRAM chips to corrupt or exfiltrate data. A common DRAM exploitation technique is a "rowhammer" attack.
- **Cache attacks:** Cache poisoning attacks involve injecting malicious data into the cache memory, enabling an attacker to execute arbitrary code or steal sensitive information. Cache denial of service (DoS) attacks overload the cache memory with requests, causing it to crash or slow down.
- **Speculative execution attacks:** Take advantage of the side effects of speculative execution to steal sensitive information. Speculative execution attacks include Meltdown and Spectre vulnerabilities created in the execution of low-level code known as "kernel code."
- **Side-channel attacks:** Target information that can be inferred from a server's physical characteristics, such as power consumption, electromagnetic emissions, or timing of the system's operations.
- **Physical attacks:** Physical tampering with the hardware, such as inserting a rogue component.

FIGURE 7
Top Priorities for IT Infrastructure

#1 **Improving security and compliance**

#2 **Enhancing agility and flexibility**

#3 **Better automation and orchestration**

Source: IDC's European Infrastructure Survey, 2022

Given the importance of IT infrastructure to organizations, security is critical, and this is reflected in IDC research showing that improving security and compliance is the number 1 datacenter priority for 2022–2023 (see Figure 7). Improving security and compliance includes taking a comprehensive and proactive approach to hardware security by incorporating security-focused design principles and features down to the core, including hardware chips and processors that store credentials and sensitive data.

The "Security by Default" Concept in Hardware Protection

Hardware security can provide the foundation for the security of a device and its data and is important to help protect a server device even before a device's software security measures are implemented. "Security by default" refers to the concept that security should be an integral and default part of the design and implementation of a system, rather than an afterthought or optional feature. The benefits of this in hardware systems are a stronger security posture, making it more difficult for attackers to exploit vulnerabilities, and improved system resilience.

To help protect against cyberattacks, organizations should consider:

- **Secure boot protection:** Helps verify the authenticity of system software, including the boot loader, operating system, and firmware being loaded and executed during hardware start-up. The feature is designed to help prevent tampering during the boot process and help detect malicious software and prevent it from executing. The secure boot process starts with a hardware root of trust (RoT), a hardware component in the device that is designed to provide a secure and trusted starting point for the boot process. The RoT verifies the authenticity of the first stage bootloader and only allows it to run if the bootloader is trusted.
- **Secure memory encryption:** Helps protect sensitive information stored in memory from unauthorized access. The encryption is performed on a per-page basis, with each page of memory being encrypted using a unique encryption key.
- **Shadow stack protection:** Used to help prevent memory-based attacks that target the call stack, the processor compares the contents of the primary call stack and the shadow

stack to see if they match. If there is any discrepancy, the processor can trigger a security exception, indicating that an attack may have occurred.

- **Application control:** Hardware-based application control that allows only approved software and executable code to run, even if the operating system or other software-based security measures have been compromised.
- **Hardware attestation:** Verification of the identity and trustworthiness of another device or system, implemented through a trusted digital certificate used to attest to a device's security posture.

Data-in-Use Protection and Confidential Computing

Sensitive and regulated data now cross multiple networks, storage systems, and geographies, all of which present potential vectors of attack. When discussing security strategies to effectively combat these threats, security experts often categorize data into its three different states:

- **Data at rest:** inactive data sets that are stored physically in any digital form (e.g., in databases, on file systems, and in archives)
- **Data in flight:** streams of data moving through any kind of network within a datacenter, on the internet, and between devices
- **Data in use:** data in computer memory or data currently being processed by applications and actively being shaped in a CPU subsystem

Data encryption use is highly recommended and considered fundamental to data security best practices. European regulators have shown they will take a tough approach to serious failings to adequately protect sensitive data. Traditional end-to-end encryption typically only protects data at rest and in transit. However, when sensitive data is needed for processing or collaboration, it is de-encrypted and transformed into a state that is unprotected and at risk of exfiltration.

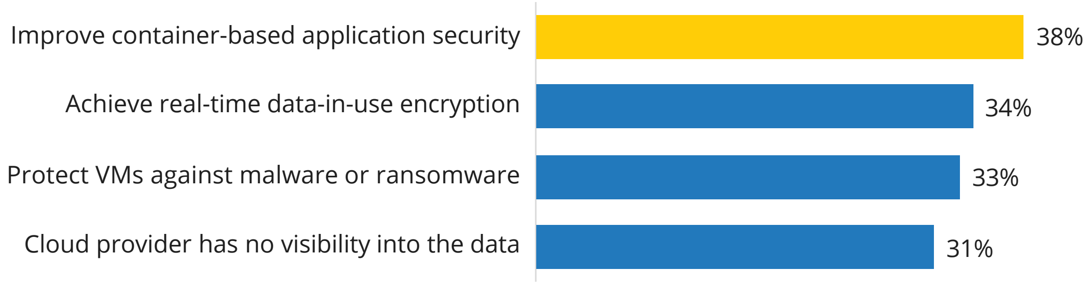
Confidential computing is an emerging technology area and represents a new paradigm in computing, aimed at addressing the growing concern around data privacy and security in the cloud and other shared computing environments. Confidential computing technologies are designed to be secure and trusted hardware-based environments that help protect sensitive data in use. Like bank vaults, confidential computing technologies allow data to be decrypted and processed in a confidential environment that can't be accessed by unauthorized parties, even by the operating system or other software running on the same machine, thereby helping to protect the data from outside environments. Confidential computing technologies can also help provide an additional layer of security if the underlying infrastructure is compromised.

Confidential computing is unlocking new compute use cases that have been difficult to implement due to privacy, security, and regulatory compliance concerns. Primarily it enables organizations with high security and confidentiality requirements to confidently move sensitive workloads off premises into the cloud, while also realizing the benefits of cloud computing. Confidential computing can also be extended to virtual machines in multitenant cloud environments where multiple customers have virtualized resources on the same host. Confidential VMs can provide significantly improved compute protection for customer data against other tenants, the underlying infrastructure, and/or cloud operators.

According to IDC research, the main reason for deploying confidential computing is to protect container workloads and VMs running in the cloud (see Figure 8). Cloud infrastructure providers such as AWS, Google Cloud, and Microsoft Azure now also offer various deployment models for confidential computing, including for virtual machines.

FIGURE 8

Main Reasons for Deploying Confidential Computing in the Cloud



Source: IDC Cloud Survey, December 2021

Confidential computing adoption is increasing and is supported by the Confidential Computing Consortium (CCC), an industry group focused on driving open standards, promoting best practices, and fostering an ecosystem of developers, hardware OEMs, and ISVs. Members include major technology companies such as AMD, Google, and Microsoft.

Security by Default: Hardware Security Features in AMD Processors

AMD is a leading provider of silicon products with a broad portfolio covering compute, graphics, and networking. AMD EPYC processors for the server and datacenter market are engineered with cutting-edge hardware-based security features collectively known as AMD Infinity Guard. This offers multilayered security features designed to provide a strong security foundation for computing in the cloud or on premises, and helps provide protection against sophisticated attacks that are difficult to stop with software alone. These attacks include BIOS manipulation, in-memory return-oriented programming (ROP), and virtualized malicious hypervisor attacks.

EPYC processors are designed to offer the following built-in security features:

- **A secure processor** creates a hardware root of trust, an embedded security checkpoint designed to validate the initial BIOS software boot without corruption from malware.
- **Secure memory encryption (SME)** helps protect against internal and physical memory attacks, such as cold boot attacks, by encrypting all system memory. SME helps encrypt memory pages as they are written to DRAM, providing a secure environment for sensitive data. Enabling SME can help prevent attackers who gain physical access to a system's memory from reading the sensitive data.
- **A shadow stack** provides a mechanism to detect and prevent code injection and ROP attacks. This provides a higher level of security for the system and helps to prevent attackers from taking control of the system.
- **Secure encrypted virtualization (SEV)** provides confidential computing capabilities for virtual machines running in a cloud environment. SEV helps with encryption of the virtual

machine memory, so even if the physical host or guest VM is compromised, the encrypted VM remains secure. Unlike other approaches, no reconfiguration of the application is required to fit the secure environment. AMD Secure Encrypted Virtualization (SEV) is available in several variants:

- SEV-ES (Encrypted State) encrypts memory and CPU register data to provide a layer of protection between the sensitive information and the hypervisor during VM interruptions and failures.
- SEV-SNP (Secure Nested Paging) builds on SEV and SEV-ES functionality by incorporating strong memory integrity protections against hypervisor-based attacks, such as data replay, memory remapping, and aliasing.

Conclusion: Invest in Server Processor Security From a Trusted Vendor

IDC recommends that organizations take a right-fit approach when investing in the next wave of IT infrastructure to support their digital-first strategies. The focus should be on a workload-optimized, innovation-focused infrastructure platform with advanced hardware security features designed in, not bolted on. Processor choice is a key consideration, as the latest generation of security chipsets enable hardware-based authentication and encryption.

Trusting an infrastructure vendor should be based on the vendor's ability to build and deliver a platform with desired security capabilities. This means the ability to:

- **Maintain a secure supply chain.** This includes verifying the authenticity of components or parts, procuring them from trusted suppliers, and physically securing the build environment, the system build process, and the process of shipping the system to the customer.
- **Harden systems.** The first step to building security into a system is to establish integrity in the foundation, reducing the attack surface and building on a root of trust.
- **Build security into every design step.** This includes incorporating proactive security features in the hardware to help prevent malicious attacks as well as developing, testing, and verifying the integrity of the firmware prior to installing it.
- **Include processor cryptographic features.** Security and integrity are built on a foundation of encryption. Cryptographic functions should include hardware-based memory encryption to protect confidential data-in-use workloads.
- **Provide proactive updates and system recovery.** This includes proactively patching firmware bugs or delivering and applying updates. When a system recovery is necessary, it provides a restore to a known good state.
- **Offer innovative security strategies.** This includes taking new approaches to security that enable IT decision makers to integrate new security features into their IT infrastructure, having to make few if any modifications to their applications.

Partnering with an IT infrastructure vendor that considers end-to-end server security a principal value proposition, offers products and solutions with built-in latest-generation hardware security, and backs up its offerings with services and support organizations is an effective approach to help protect, detect, and recover from security threats.

MESSAGE FROM THE SPONSOR

For more than 50 years AMD has driven innovation in high-performance computing, graphics, and visualization technologies. Billions of people, leading Fortune 500 businesses, and cutting-edge scientific research institutions around the world rely on AMD technology daily to improve how they live, work, and play. 4th Gen AMD EPYC is ideally placed to help enterprises gain a competitive edge. Datacenter computing powered by AMD EPYC processors delivers leadership, application performance, TCO, and security features to help your business scale.

Find out more about AMD Data Center Solutions at amd.com/en/processors/epyc-9004-series.

About the Analyst

Ralf Helkenberg, Research Manager, European Privacy and Data Security, IDC

Ralf Helkenberg is a research manager with the European security research team. He leads IDC's European Privacy and Data Security research practice, and his core research coverage includes the impact of data protection regulations such as GDPR on the technology sector, with key insight into market dynamics, vendor activities in privacy workflow management and data security (including data discovery, DLP, encryption), end-user trends, and the future of digital 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.

With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives.

Founded in 1964, IDC is a wholly-owned subsidiary of International Data Group (IDG, Inc.), the world's leading tech media, data and marketing services company.

IDC UK

5th Floor, Ealing Cross,
85 Uxbridge Road
London
W5 5TH, United Kingdom
44.208.987.7100
Twitter: @IDC
idc-community.com
www.uk.idc.com

Global Headquarters

140 Kendrick Street,
Building B
Needham,
MA 02494
+1.508.872.8200
www.idc.com

IDC Custom Solutions

This publication was produced by IDC Custom Solutions. As a premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets, IDC's Custom Solutions group helps clients plan, market, sell and succeed in the global marketplace. We create actionable market intelligence and influential content marketing programs that yield measurable results.