Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

I D C E X E C U T I V E B R I E F

The Trusted Computing Platform Emerges as Industry's First Comprehensive Approach to IT Security

February 2006

By Shane Rau, Program Manager

Sponsored by Trusted Computing Group

Introduction

The external threat to computing devices is escalating. Even as threats like worms and viruses multiply, the number of vulnerable devices proliferates, and wireless technologies provide more avenues of attack. These forces are driving organizations to seek more fundamental and broad-based security solutions that ensure privacy, data integrity, user authentication, and access control at the server, the client, and across the network. This Brief examines the nature of today's escalating information-security threats to computing devices and describes the efforts of the Trusted Computing Group to build security into computing devices.

The Information Security Threat Profile

From worms and viruses to hackers and phishers, the multiplying threats to computer security forces examination of their nature and scale and what's being done to protect privacy, data, and system manageability from them.

Threats compromise security in one of three ways. They can compromise the confidentiality of data by violating privacy, the integrity of data by altering or destroying it, or the usability of the environment around the data, which means attacking how the system or network works or is structured.

Due to their potential to compromise security in all three ways, viruses and worms are still considered the most serious threat to corporations. However, spyware has rapidly climbed the priority list, and now ranks as the second-most serious threat, partially due to its violation of privacy, but primarily due to its scale and consumption of system and network resources. IDC believes that spyware infects more than three-quarters of all corporate machines, which implies significant theft of confidential information, loss of employee

productivity, consumption of bandwidth, damage to corporate desktops, and a spike in the number help-desk calls.

Spam has climbed back up the priority list of IT managers and security departments and ranks as the third-greatest threat to enterprise security. Similar to spyware in the primary nature of its threat, spam continues to increase rapidly and consume resources by clogging networks, servers, and inboxes with unwanted and often offensive content. In fact, most spam is now automated, being sent by a botnet of zombie machines remotely controlled by spammers. As a result of the extremely rapid growth in the volume of spam, the convenience and efficiency of email have been dramatically reduced.

Phishing, a method of identity theft, has surged recently and attacks are now daily occurrences for any organization, but especially for the largest financial institutions and their customers. Large Internet service providers (ISPs), security technology providers, and law enforcement agencies are paying special attention to these attacks, which can be performed through email, where victims are asked to provide personal information such as passwords, credit card numbers, or account information in a reply email. The criminals then use the personal information for their own purposes or sell the information to other criminal parties.

Beyond the number and scale of threats, proliferation of computing devices and the methods of connecting to them are aggravating the collective threat.

Cross-Platform Threats

Handheld devices, including mobile phones, personal digital assistants (PDAs), and converged devices (a.k.a. smart phones), have grown in sophistication to the point where they are also legitimate computing devices. They store large quantities of data, have operating systems, support user IDs, and are increasingly integrated into corporate networks, if only as email platforms — which means they can offer entrée into those networks. As a result, IDC believes that handheld devices will be targets of threats similar to those that afflict traditional PCs.

To date, several viruses have been specifically developed to exploit the vulnerabilities of handheld devices and, while the majority of these have been low-level threats, they have laid the "proof-of-concept" groundwork for others to follow. As newer, third-generation (3G) phones and handheld computers become more powerful, they will be more susceptible to the techniques used by many desktop computer viruses. Also, the fact that mobile-device networks are typically always-on network connections for email and text messaging, they allow handheld-device viruses to spread much faster than traditional viruses.

IDC also expects other malicious efforts, such as spam, spyware, and phishing attacks to quickly move to the mobile phone platform. Just as PC users struggle with spam filling their inboxes, unwanted text messages are just starting to become a growing nuisance in the mobile world. Mobile spam can be more costly, too (e.g., some unsolicited texts are designed to trick users into phoning premium rate numbers).

Traditional hacking techniques are also making their way to mobile devices, with phishing and denial-of-service attacks possible due to Bluetooth hacking tools that are freely available on the Internet. Hackers can trick unsuspecting mobile users into turning on Bluetooth, for example, by simply sending a message and having the user accept it. The hacker can then link the two devices together and gain access to the user's mobile device — a direct threat to both personal and corporate data stored on mobile devices.

The proliferation of computing devices has also meant a proliferation of connections. Mobile PCs — shipments of which grew 30% in 2005 — have gone wireless but, still needing to connect to the traditional wired infrastructure wherever they may be, have access points distributed widely. Furthermore, as the aforementioned handheld devices have emerged as legitimate computing platforms, both PCs and handheld devices are developing the infrastructure to connect to each other's networks.

IDC estimates that not only will 100% of mobile PCs ship with WiFi technology by 2009, but nearly 5% of them will also ship with wireless WAN (cellular) technology in that year. IDC believes it is only a matter of time before the wireless world is hit by the same sorts of malicious viruses and worms that attack corporate networks and desktops.

Requirements

The number, scale, and breadth of threats to computing device security are driving organizations to develop new ways to protect the confidentiality of data, the integrity of data, and the usability of the environment around the data. Unfortunately to date, security solutions such as smart cards, antivirus software, and firewalls have been peripheral to system and network design. Like putting a moat around an ordinary house, the entire security environment would be better if the house were more like a castle, built with security fundamental to its purpose.

Making security fundamental requires defining the terms of a "healthy" computing environment for businesses and individual users. In this way, variations from what is healthy thus move into the class of unhealthy. A healthy environment demands trusted users, data, systems, and networks that are all capable of proving trustworthy by what they are and how they behave in the environment. Terms used to define these levels of trust are authentication and access control.

Solution and Benefits

The computing industry, responding to organizations seeking more comprehensive and integrated security solutions, has started down the road to make security fundamental to computing devices, thus creating a healthy computing environment.

The industry's first attempt to do so has emerged in the form of the Trusted Computing Group (TCG). Formed in 2003 by several major IT providers, TCG succeeded the Trusted Computing Platform Alliance. TCG members include major silicon providers, system OEMs, ISVs, and IHVs. TCG's mission is to develop standards-based security solutions that mitigate the risks of participating in an interconnected world while also ensuring interoperability and protecting privacy.

TCG's approach toward achieving Trusted Computing begins with setting out specifications that are open for all its members to share in common. Key to its specifications is that security should literally be part of devices. The first application of Trusted Computing in PCs, for example, makes security a part of the system in the form of a semiconductor chip built into the computer motherboard. Dubbed the Trusted Platform Module (TPM), this chip (technically a microcontroller) stores passwords and digital keys that represent the PC's unique identity.

TPMs can either be separate (discrete) chips or integrated into other chips, such as Ethernet controllers. As part of the system that tracks what is a healthy state of the system, the TPM becomes the root of trust. It can benefit users and their systems by securing the PC itself, authenticating the trusted users are who they claim to be and controlling access to data and networks, thus lowering the threat from outside attacks and from loss of the system itself.

Like a foundation of a building, the TPM supports all security-related components layered over it. Immediately above the TPM in this virtual stack is TCG Software Stack (TSS), a driver that provides a standardized method for software applications, services, and hardware devices higher in the stack to access the TPM and take advantage of its capabilities. For example, an encryption application could access the TPM for its key.

The TSS could enable ISVs, IHVs, and service providers to develop security-related solutions on a common framework, and so avoid proprietary solutions that can inhibit widespread development.

A use case of the TPM and the overall TCG security infrastructure would be a banker who needs access to critical financial files on a bank's server through her corporate TPM-enabled desktop PC. Upon booting, her PC's operating system checks with the TPM to confirm that the system's configuration hasn't been altered since her last session. After this self-authentication, the system then prompts the banker for her smart card and password, checking her response against the password stored in the TPM and on the smart card.

When she tries to access the critical files on the server, the server authenticates her PC's identity and confirms that she and her PC are entitled to access the files. These files, encrypted on the server and while in transit over the network, are not decrypted until they reach her system and the system's TPM has provided the necessary key for decryption.

Beyond PCs, TCG is extending the building blocks of Trusted Computing to networks and other device types. Implicit in the use case is security inherent not only in the PC used by our banker, but also in the bank's network and server. TCG's Trusted Network Connect (TNC) specification provides for authentication and access control at the endpoints between a client and a server.

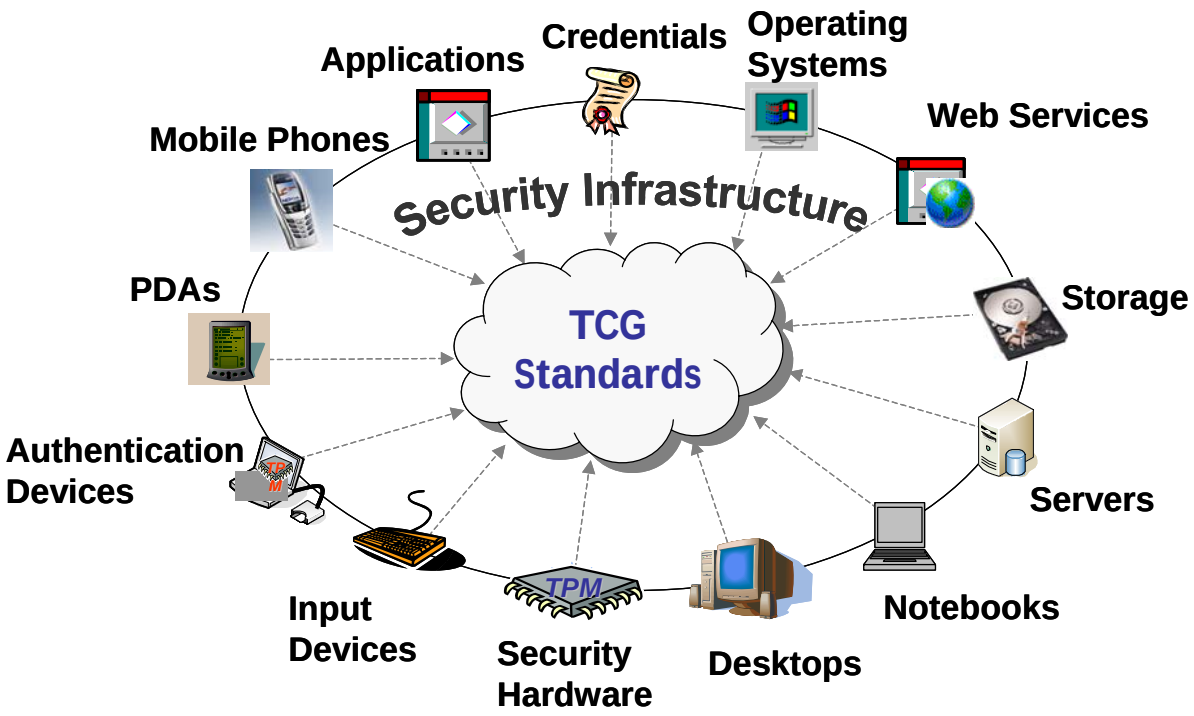
The need for the TNC acknowledges, even if a TPM authenticates a user on a particular PC, that the user or their PC may not be trusted any further, particularly if beyond their PC's network connection lie the PCs of many other users and the server. TCG has also developed the Trusted Server Specification which anticipates that, even though servers can technically use the same TPMs and software stacks developed for PC clients, the user load on servers will likely require more robust, higher bandwidth versions of these building blocks.

TCG also is extending Trusted Computing to handheld devices and to storage devices. While seemingly far afield from traditional PCs, servers, networks, and each other, handheld and storage devices are increasingly intelligent clients seeking wired and wireless access to traditional computing networks, such as the WLAN and WWAN networks identified earlier. Thus, they too, will need to opt-in and provide mechanisms of authentication and access control that define their healthy environments and prove their trustworthiness to other environments.

Figure 1 illustrates how broadly TCG intends to extend its specifications and security technology building blocks.

Figure 1

Trusted Computing Security Ecosystem



Source: TCG, 2006

Key Market Trends

TCG's approach to embedding security reflects a long-term IT market trend that IDC dubs "the platform approach." This comprehensive approach to system design considers all the components of that system and how they will be used. TCG's multitiered approach — incorporating hardware, software, services, and the connections between platforms that need to remain secure, — involves multiple suppliers in the supply chain of a device into the foundation of the devices' security. It's also about secure solutions delivered by the ecosystem through these TCG building blocks and specifications.

At the silicon level, several vendors produce TPMs. While most produce discrete versions, at least one vendor integrates the technology into a gigabit Ethernet controller, and another has integrated the TPM into an I/O chip.

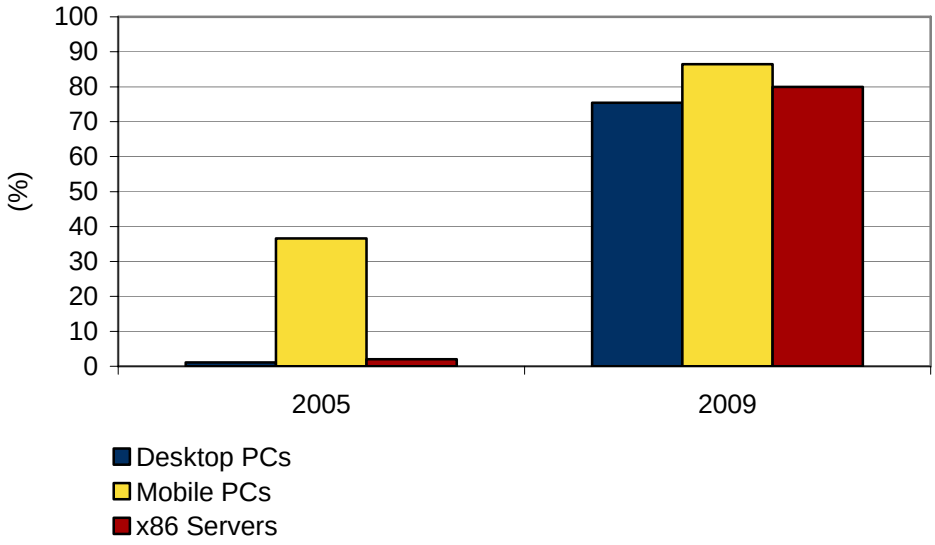
At the system level, notable adoption of TPMs began in 2004 when leading vendors began adopting the IC in some of its PCs. Today, most leading PC makers are incorporating the TPM into one or more of their computer models. IDC expects that most TPMs will be

integrated because they can be offered as a value-added feature for little or no additional cost and also minimize the space they occupy. However, we also believe that, due to certification requirements that lend themselves to a discrete IC, a notable number of corporate servers and desktop PCs will retain discrete TPMs.

Figure 2 compares TPM adoption by form factor in 2005 and 2009. Underlying our forecast is the assumption that the next major factor in TPM adoption in desktop PCs, mobile PCs, and PC servers will be support for TPMs within the leading desktop operating system. IDC believes that TPM adoption by computing OEMs will accelerate and stimulate development of the software infrastructure — drivers, applications, operating systems — that will enable more effective TPM-based security.

Figure 2

Worldwide Trusted Platform Module PC Penetration Forecast by PC Form Factor



Source: IDC, 2006

Outside of PCs, IDC believes that ecosystem suppliers for service, network, mobile, and storage devices will also look to integrate the building blocks of security commensurate with the usage model of their end devices. A major hard drive manufacturer, for example, has introduced a mobile PC hard drive that integrates features suitable for mobile PC usage, including full disk encryption. In this way, the company presages how security can be good not only for its own sake, but also as a way to drive higher margins across the ecosystem.

Considerations

A major challenge facing TPM adoption lies in the fundamental nature and sheer scale of the effort. As an attempt to embed security in all hardware, software, and networked levels of the computing device, TCG is a monumental effort.

Another challenge facing TPM is that the system is only as good as far as it is used. Even assuming that Trusted Computing technology becomes ubiquitous, it must be turned on by the individuals and enterprises that purchase it. Furthermore, given its design as part of several hardware, software, networked layers, it cannot be utilized in isolation. Other hardware and software measures, including smart cards, virus monitors, and firewalls, while individually more vulnerable, are essential to the overall goal of security.

Lastly, TPM is only as good as the capabilities of the people using the system. The technology cannot anticipate what uninformed or naïve people will do to compromise security. Social engineering attacks, including phishing, in which an individual is tricked into giving up access or secure information, can only be prevented by changes in the human element, which is likely the weakest link in all security arrangements. Education around computing, networking, the Internet, and security is essential for users in enterprise and individual consumers alike.

Technological fixes, whether through improved standards, new products, or integrated solutions are not sufficient for enterprise security without informed users. Working with users to ensure they are using security products properly is just as important as working with vendors to make more secure products.

Conclusion

The approach to security in computing devices is a matter of all stick and no carrot. A growing number of profit-minded perpetrators and increasingly sophisticated attacks speak for themselves as threats, and the proliferation of wired and wireless networking and the devices connected to them speaks to the growing threat matrix. If such a threat profile is not sufficient to boost concern among user organizations, then corporate deadlines for compliance with government and industry regulations should suffice.

Comprehensive security requires solutions at the user level, the system level, and the network level. The approach of the Trusted Computing Group is the first to attempt a comprehensive, platform-based security solution that ensures privacy, data integrity, and user authentication at all these levels.