Carnegie Mellon CyLab

Information Security for the Next Century

Why we need an information-centric approach to data protection

Dr. Pradeep Khosla

Contents

Today's data protection challenges

Data breaches are becoming daily news, where organizations put personal information about millions of consumers at risk by not taking adequate measures to ensure its protection. The results? Violations of personal privacy, bank account and credit theft, misappropriation of competitive intelligence, and compromised national security. Breaches are affecting everyone, the costs for cleaning them up are skyrocketing, and the impact on the average person is tangible.

In fact, a recent report from the Identity Theft Resource Center concludes that the number of breaches in 2008 grew by 47% from the previous year. The cost of a breach has also risen from an average of $6.3M per breach to $6.6M over the same period, according to the latest Ponemon *Cost of a Data Breach* study (Ponemon, 2009). The Privacy Rights Clearinghouse, which runs a public Web site chronicling data breaches from 2005 on, details domestic incidents involving a total of 244,803,916 compromised records as of late summer of 2008 [http://www.privacyrights.org/ar/ChronDataBreaches.htm]. New cases are being logged into the PRCH database every day, demonstrating the colossal liability of organizations that hold sensitive personal data.

The need to comply with state, federal, and industry regulations is driving increases in data security deployments, and those regulations are becoming continuously more stringent. Forty-four states have passed breach notification laws that can impose punitive damages or open businesses up for class action law suits. Massachusetts has the strictest law to-date, requiring companies to encrypt all mobile personal data, which includes data on a bevy of ubiquitous mobile devices.

With such massive financial losses facing organizations, why are the massive holes in enterprise security not being filled more quickly? Is the security industry failing to build effective solutions to address the problems?

The Problem

Many of the answers to these questions can be traced to organizational issues, technical challenges, and the fact that companies may not address the right security threats. At the heart of the problem is the real nature of information security and its goals within the organization.

Too often, security is associated with ***restricting the sharing*** of data. This is the most common of all security myths. In fact, security should enable confidence in the sharing of data with those who need it. In this information age, flows of information among organizations, partners, and employees are the core foundation for innovation and increased productivity. In the case of national security, this sharing and collaboration of information among the various agencies can be vital in making the right decisions. If security becomes a hindrance to the free movement of information to critical constituents, it will always be ignored or relegated to a less important role. Security can be even more problematic if it hinders the flow of information and result in sub-optimal decisions.

The main cause of insider breaches is an employee "working-around" a security process or policy. Restrictive and cumbersome security processes are less effective because users fight their adoption and the adoption of respective security technologies. The right approach is to ensure optimal security while enabling free information flow.

A second challenge is the role of security in an organization. A National Cyber Security Alliance study found that over 73% of business users believed they had nothing to do with ensuring the security of data — their IT department should be dealing with it. This couldn't be further from the truth. CISOs (Chief Information Security Officers) may have responsibility for data security; however, many don't have the authority to enforce it. This lack of ownership from the business for security hinders the best security solutions and processes from benefiting the organization.

Information security is a huge organizational challenge that requires involvement from the very top of the organization since it involves business processes, workflows, and user behavior. In fact, if security is not designed with user workflows in mind, the result would be reduced productivity and frustration among the end users. Technology can only be the foundation on which good security is built.

The current security environment

With the explosion of devices, the ubiquity of networks, and the distributed nature of business, organizations find it increasingly challenging to protect sensitive data. They are burdened with multiple point solutions that only secure a specific device or network link. The focus of today's solutions has been in protecting (encrypting) devices in order to protect data resting on them, or protecting the networks in order to protect the data passing through them. In short, the data protection is mostly protection by proxy.

These point products have to be integrated to provide a more complete security capability, which makes the infrastructure complex and costly to maintain. In addition, transition points among these products become the weak link in the security chain because they are vulnerable to breach. These solutions fail to address the bigger problems of information sharing and collaboration since they are so focused on protecting the infrastructure. Many solutions are also difficult to work with because they require changes to applications or user behavior.

In response to the growing threats (and in spite of the sub-optimal nature of solutions), enterprises all over the world are increasing their already-large investments in this device-centered data security. Yet the frequency of data breaches continues to grow and the collateral damage continues to spread.

Even though the value of the data in the breach examples cited here is genuinely high, it actually pales next to future scenarios. If we extrapolate current trends in data storage, by 2020 a desktop computer will be able to store — and potentially to share — files measured in petabytes, the magnitude of which could easily hold the digital representation of an entire human brain. As that capacity grows, so will the value of the information it holds. But instead of running such sensitive data in locked-down data centers, the software will almost certainly be executed on third-party or public processors distributed across the globe to take advantage of enormous distributed computing power.

Fixing the problem: An information-centric approach

With the proliferation of computers, devices, servers, ports, and networks over which data can be distributed, it is almost impossible to control each device to prevent unauthorized access. Most products that secure enterprise data are focused on specific points, like a laptop, port, USB device, or network. But as distributed data explodes across the enterprise, those points are multiplying so fast that traditional solutions can't keep up. Data needs to be secure and managed everywhere it goes, as it moves and while it is stored. But with the multitude of devices and networks in the modern IT environment, that hasn't been possible without burying IT staff in a mass of different technologies that is both too complex to manage and too costly to maintain.

Despite the varied nature of these devices, networks, and applications, one factor remains constant: the data itself. Continuously protecting the data itself, regardless of which device it is on or network over which is travels, is the only way to ensure that protection will be ubiquitous and scalable.

A better approach to data protection

By thinking about data security in a different way, not only can we rethink the goals of data security, we can overcome many of the challenges facing IT organizations. If security technology can't facilitate easy sharing of information, even with external partners and suppliers, it ultimately affects the bottom line. In today's business environment, that simply isn't acceptable.

The only way to ensure end-to-end data protection while enabling secure collaboration is to protect the data itself; as opposed to protecting the devices that store the data or the networks that data passes through. Today's data protection solutions are mainly focused on protecting devices or networks; i.e. infrastructure protection similar to the traditional "moat and castle" model. In light of amazing increases in the mobility, distribution, and value of data, more and more IT managers are questioning the viability of only using secure perimeters around stationary computing devices. The pain is particularly acute because, as the number and size of data files keeps growing, the number of devices and pathways which store and transmit their information will soar exponentially. In short, the more you have, the more you have to protect. In today's distributed world, a separate method of protection for every device and network is ineffective and unacceptably complex.

The information-centric approach to security

The logical alternative is an information-centric approach to data security, which shifts protection from devices or networks directly to the data itself. With information centricity, data can be persistently protected and remains so at rest and in motion, at all times, wherever it goes. It allows for device and network independence since the data protection follows a piece of data regardless of where it is. The information centric model also means that security policies are applied at data level — each data element contains policies that explain the rights of users and the actions that can be taken on that data — and these policies are embedded with the data itself, making data self-protecting.

Benefits of information-centric security

There are significant advantages to the information-centric security approach.

- **Protect data wherever it goes**. The main benefit of an information-centric approach is the ability to protect the critical asset at all times, without interruption. Unlike perimeter defenses, it protects data from both insider and outsider threats and has the added benefits of device independence, network independence, and the ability to protect data in virtual environments. The data becomes smarter and self-defending, and is therefore much easier to share and collaborate with.
- **Align with business flows.** Since policies for data use are always with their respective data, that data can be shared and collaborated on with more confidence. Legitimate users are not restricted to certain devices or networks since the appropriate security and access policies will ensure and enforce user rights regardless of where the data is. Quite possibly the biggest advantage of this approach, however, is that it allows business users — not the IT department — to truly own information, the rights associated with it, and its flow across the organization. Business can take responsibility for the security of data, since they feel they own the data itself. IT's primary role will be to focus on computing devices, servers, networks, etc. Ultimately,

information-centric data security can bring about better partnerships between business and IT since the lines of responsibility for security can be clearly drawn.

- **Reduce costs and complexity.** The device independence of information-centric security significantly reduces the number of separate device-centric or network-centric security solutions. Subsequently, it lowers acquisition costs, maintenance burden, and man-hours associated with integrating multiple pieces of hardware and software. The information-centric approach protects critical data assets themselves, regardless of the device or network that carries them. An organization can secure data with far fewer solutions and with far less man power.
- **Increase end-user transparency.** A major cause of data breaches is when legitimate users, while trying to be productive, work around security restrictions. Why would they do such a thing? Because following the security practices dictated is often inconvenient and creates more work for them. A security solution should remain as transparent as possible to end users. If user workflow is not hindered or altered, there is a significantly higher chance that the security program will be followed, and hence be more effective. Information-centric security can be extremely transparent, since the protection is with the data itself. Users do not have to explicitly make decisions about valid devices, network authentications — all these policies are contained in the data itself and can be managed centrally, so they can operate as usual without interference.

The requirements of an information-centric security platform

The basics of information-centric security exist today, and they're very well suited to commercial and federal data protection scenarios, particularly those answering the protection requirements of regulatory compliance. The following are the basic requirements for a true information-centric approach to data protection.

1. Smart data with embedded policies

Before data can be adequately protected, it needs to be easily identified and managed; you can't hit a security target if you don't know where it is or what it looks like. If data objects are enriched with metadata tags that carry security policies, that data can be empowered to protect, replicate, or even delete itself, as required. As a result, an information-centric security approach could enable files to communicate their vital characteristics to the devices they pass through, as well as to other data objects, throughout the cycle of their lives.

2. Universal policy language

To realize truly effective and universal information-centric security, security policies and the codes describing them need to become industry-wide standards. As data travels between servers, laptops, and removable media, the policies that govern its protection need to be enforced, regardless of the platform. To make sure those policies are interpreted uniformly throughout an enterprise, there has to be a common policy language, embedded in the data itself, that every device can understand. That way, no matter where the data moves or resides, or the security solution that protects the data, the policies governing its access remain in force.

3. User-friendly implementations

Any successful information-centric security solution needs to be transparent. That is, users shouldn't have to modify their work habits or change their business practices in order to benefit from the security solution. If they do, the solution will almost certainly fail. Why? Because most people will reject changes imposed on their familiar work patterns and bypass new security provisions they consider a nuisance. Nor should an organization's current software applications or computer platforms have to be upgraded as part of the security deployment; an information-centric approach is capable of working with any device, on any platform, without requiring special patches or programming. Otherwise, implementation costs will be prohibitive. Beyond that, essentially all IT environments today utilize legacy systems to some extent. That makes it difficult for administrators to justify major overhauls solely for the sake of security. But by taking an information-centric versus a device-centric approach, it is possible to create a security solution that can apply to multiple IT environments and, at the same time, avoid having to inconvenience users.

Future trends and applications

Enhancing Data Leakage Prevention Solutions

DLP solutions discover and classify data based on their content and other criteria to decide their importance. They discover data on the network, end points, email gateways, file shares, etc., and they use policies to prevent data from leaving the enterprise by blocking networks, ports on laptops, etc. However, most of the solutions cannot persistently encrypt or protect that discovered data with appropriate access controls or classification policies. In short, since the DLP solutions do not take preventive and protective action on data, they fail to complete the full cycle of protection.

Information-centric security can complement a DLP solution by persistently protecting sensitive data with appropriate policies for encryption, access control and classification. Its capability to embed policies with the data itself ensures protection of data at rest on various devices or in motion across the network. By focusing on the data itself, information-centric approaches reduce the multiple integration points needed for DLP solutions to achieve better leakage prevention.

Virtualization security

Any computing environment requires four elements: devices/OS, networks, applications, and data. With the advent of virtualization, physical devices are being replaced by flexible, on-demand virtual "devices," networks are being virtualized and applications are being streamed down from virtual environments. Therefore, the only remaining "constant" element is ***the data itself***. In addition, data has a longer lifetime and is thus more vulnerable than the virtual environment itself which is created and brought down based on business and operational requirements.

However, virtualization security solutions today primarily focus on protecting the virtual OS, the virtual networks, or the hypervisor software itself. That is, the focus is mainly on protecting infrastructure and perimeter, not data. While protecting the virtual infrastructure is important, the primary focus for protection should be the data — the true IT asset. Virtualization is a game-changer for computing and has forced the IT world to rethink its infrastructure; now virtualization security has to be rethought as well.

This information-centric security approach ensures that security policies that travel with the data, at rest and in flight across multiple devices, networks, and virtual environments (including mixed virtual and physical environments), thus enforcing protection ***wherever the data is***.

The current approach of porting security solutions from the physical world will stifle the promise of virtualization. An information-centric security approach is the only way to truly take advantage of the real benefits that virtualization can bring about.

Summary

Information-centric security systems will inevitably become the standard for protecting sensitive data for several basic reasons:

- Current device-centered solutions work only in limited environments, and those environments are becoming increasingly irrelevant to most users
- Information-centric protection offers a more efficient, more cost-effective, and more scalable approach to safeguarding sensitive information
- Collaboration and ensuring confidence in sharing sensitive data with authorized users will require protection policies to be embedded with data
- Enable better collaboration between IT departments and business units
- Technology is quickly advancing to the place where information-centric solutions are viable.

With the rise of virtualization, information-centric security will gain even more importance simply because data will be the only permanent, stable element of an IT network. In a virtualized environment, there won't be anything else to protect! Current security solutions that take advantage of information-centric architecture lay the foundation for better protection, and as infrastructure continues to evolve, they'll be the only way to achieve data security.

Information-centric security solutions will inevitably become the standard for protecting sensitive data and aligns closely with how organizations collaborate and work. Not only do they offer a more efficient, more cost-effective, and more scalable approach to safeguarding sensitive information, they are the manner by which data will be protected as we move into the era of virtualized computing environments.