Data Loss Prevention

Introduction

In today’s digital economy, data enters and leaves enterprises’ cyberspace at record rates. For a typical enterprise, millions of emails are sent and received and thousands of files are downloaded, saved or transferred via various channels or devices on a daily basis. Meanwhile, enterprises hold sensitive data that customers, business partners, regulators, and shareholders expect them to protect. Unfortunately, companies constantly fall victim to massive data loss and high profile data leakage involving sensitive personal and corporate data continue (http://opensecurityfoundation.org/). Data loss could substantially harm a company's competitiveness and reputation, and could also invite lawsuits or regulatory crackdown for lax security. Therefore, organizations should take measures to understand the sensitive data they hold, how it is controlled, and how to prevent it from being leaked.

Data Loss Problem

According to the Open Security Foundation, which tracks publicly reported incidents, 714 cases of data loss were reported in 2008, affecting a total of more than 86 million records [1]. Depending on the type of data loss experienced, an organization can suffer a variety of consequences, but in nearly all cases there is both a financial and reputation cost.

Types of loss

Data loss can be divided into two sometimes overlapping categories:

- **Leakage –** in which sensitive data is no longer under the control of the organization. (In computer security parlance, this is a loss of *confidentiality*.) The most common form of data loss, leakage often results from hacked customer databases, and its most common consequence is potential identity theft. In the largest single attack of this type to date, 130 million credit card records were stolen from one of the US’s largest payment processors. Another involved 94 million customer records held at a major retailer.

- **Disappearance or damage –** where a correct copy of the data is no longer available to the organization (corresponding to a compromise of *integrity* or *availability*). An example of this occurred in 2009 when a major cellular phone service provider suffered widespread loss of customer data that was supposed to be housed on a third-party cloud-based storage service. In normal operation, on power-off the smart phone would automatically sync its data with the central server, which stores it to be available when the phone is used again. For reasons that are still not fully explained, a server crash at the storage service wiped out backups of memos, photos, and other data for more than a million smart phone customers.

Clearly, if the last accurate copy of data is physically stolen, the organization is facing both problems. Furthermore, in some cases it may not be clear which of these situations obtains. For example, one of the most common problems is theft or loss of laptops that an employee had taken out of the office. If the employee was updating and editing information on the laptop using multiple data sources, that copy may be the most current, and it may not be clear how to reconstruct the correct version from backup files. Without assurance of current and accurate records management, the firm may not be able to determine which records are affected.

Consequences of loss

As with other security incidents, data loss incidents can result in significant cost, but the duration and magnitude of costs vary with the type of data loss. Financial records can usually be reconstructed, and fraud incurred may result in no loss to the customer if national laws require that financial institutions bear this cost. Costs to the organization may be much more severe and may include liability costs that are not always covered by corporate insurance policies [2].

While loss of payment processing data may require years to repair, consumers generally are able to clear up problems and recover financial losses. Today's movement toward extensive use of electronic medical records can, however, present a new class of risk for both the consumer and the organization. For these records, the risk is to privacy, so if records become public, the damage to the individual is permanent rather than temporary as with a fraudulent credit card charge. Consequently, the organization may face increased litigation or regulatory penalties.

Why Data Loss Prevention

Key drivers of establishing data loss prevention mechanisms include regulatory compliance and intellectual property protection.

Regulatory Compliance

Today, many companies fall under oversight of government and industry regulations that mandate controls over information in general and personal identifiable information in particular. Major US regulatory mandates include the following, and most nations have similarly strong rules:

- Health Insurance Portability and Accountability Act of 1996. It required that to ensure privacy and confidentiality all patient healthcare information be protected when electronically stored, maintained, or transmitted.
- Gramm-Leach-Bliley Act of 1999. It mandated privacy and the protection of customer records maintained by financial institutions.
- Privacy Act of 1974. It prohibited disclosure of information in personal records by any means of communication to any person or to another agency, except pursuant to a written

request by, or with the prior written consent of, the individual to whom the record pertains.

- Federal Information Security Management Act of 2002. It provided a comprehensive framework for ensuring the effectiveness of information security controls over information resources that supported Federal operations and assets.
- Payment Card Industry Data Security Standards. A worldwide information security standard created to help organizations that process card payments to prevent credit card fraud through increased controls around data and its exposure to compromise.

Government and industry regulations are arguably the biggest drivers of data loss prevention. In addition, many states have passed data privacy or breach notification laws that require organizations to notify consumers when their information may have been exposed.

Intellectual Property Protection

According to the World Intellectual Property Organization (http://www.wipo.int/portal/index.html.en) intellectual property is creations of the mind - inventions, literary and artistic works, symbols, names, images, and designs used in commerce. For many enterprises, intellectual property may be more valuable than its physical assets. As a result, establishing policies and mechanisms to guarding against intellectual property loss or theft is critical to protect brand and maintain competitiveness of many enterprises.

Data Loss Prevention Approach

Data loss prevention is an enterprise program targeted on stopping various sensitive data from leaving the private confines of the corporation. With the recent high profile data loss incidents in the industry, data loss prevention technologies are emerging as important information security and privacy controls.

Loss Vectors

Enterprise data generally exists in the following three major states.

- Data at rest – Data residing in files systems, distributed desktops and large, centralized data stores, databases, or other storage methods.
- Data at the endpoint – Data residing at the endpoints of the network such as laptops, USB devices, external drives, CD/DVDs, archived tapes, MP3 players, iPhones, or other highly-mobile devices.
- Data in motion – Data moving through the network to the outside via email, instant messaging, peer-to-peer (P2P), File Transfer Protocol (FTP), or other communication mechanisms.

Data in each state often requires different techniques to loss prevention. For example, while deep content inspection is useful for data in motion, but not so much for data at rest. Therefore, an effective data loss prevention program should adopt appropriate techniques to cover all the loss vectors an organization has the potential to encounter.

Solution Components

An effective data loss prevention program should consist of the following essential components:

- **Manage –** Define enterprise data usage policies, report data loss incidents, and establish incident response capability to enable corrective actions to remediate violations. Data loss prevention is not just a technology issue; it is also a policy and policy management issue. Enterprise data usage policies should address issues such as how access to data is determined; how data access is authenticated; and how policies are enforced. Management functionalities should also include data loss reporting capability and incident remediation workflow management**.**

- **Discover –** Define the sensitivity of enterprise data, create an inventory of sensitive data, locate sensitive data wherever it is stored, and manage data cleanup. This includes discovering and inventorying sensitive data at rest in file servers, databases, document and records management, email repositories, and web content and applications; and scanning for sensitive data stored on the endpoint including laptops, desktops, and workstations at remote offices in order to inventory, secure, or relocate that data.

- **Monitor –** Monitor the use of sensitive data, understand sensitive data usage pattern, and gain enterprise visibility. This could include monitoring data in motion by inspecting network communications such as email, Instant Messaging (IM), web, FTP, P2P and others for confidential data in violation of data security policy; and monitoring data at the end points such as downloading to local drives, coping to USB or other removable media devices, burning to CD/DVDs, and printing or faxing electronically.

- **Protect -** Enforce security policies to proactively secure data and prevent sensitive data from leaving an enterprise. Automatic protection of sensitive data across endpoint, network and storage systems. This includes protecting data at rest with automatic encryption, quarantine, and remove; Restrict printing, saving, copying, accessing, movement of and downloading of sensitive data to removable media or other drives; and stopping data in motion from being sent in violation of data security policy or encrypting data for secure exchange.

When properly integrated, these four essential components offer effective protection of enterprise valuable information assets.

Best practices

Prioritize loss vectors. Data loss prevention is a complex problem. While a comprehensive program to address all relevant aspects of data loss is the ultimate goal, it makes far more tactical and financial sense to begin by protecting the data that represents the most danger to an enterprise. This means first identifying all the potential data loss vectors in the organization and then prioritizing them based on criteria such as past breaches, volume of communications, volume of data, the likelihood of a breach and the number of users with access to those vectors. Focusing first on the most significant and highest impact areas makes it easier to justify solutions and get started on plugging the leaks.

Protect without disruption. Data loss prevention solution should not interrupt legitimate business activities. To work effectively, a data loss prevention solution must operate without diminishing system performance or preventing workers from doing their jobs. Solutions that do not scale can cause performance issue as companies grow. Solutions that are not properly tested and tuned can also cause both false positives and false negatives that drain valuable resources.

Flexible & modular architecture. Solutions for data loss prevention are still evolving, with no single one providing all capabilities most organizations require. Enterprises need to address the data loss problem by creating a flexible and modular architecture. A flexible and modular architecture allows enterprises to immediately and cost effectively addresses their most pressing requirements while being able to add new controls as their needs change. It also ensures speedy deployment, protect investments, and easily scale to accommodate expansion and growth.

Conclusion

Data loss prevention is a serious challenge for companies, as the number of incidents continues to increase. Data loss can diminish a company's brand, reduce shareholder value, and damage the company's goodwill and reputation. Data loss prevention is not a silver bullet. Identifying and blocking all sensitive data is neither possible as an outcome nor wise as a goal. However with a more focused goal of preventing the most damaging leaks and establishing better ways for users to exchange information securely, data loss prevention can be effective, practical, and successful.

References
[1] http://datalossdb.org/yearly_reports/dataloss-2008.pdf

[2] Identity Theft: Is It Covered? American Bankers Association, *Safe Talk*, October 2007.

We identify certain products in this document, but such identification doesn't imply recommendation by the US National Institute of Standards and Technology or other agencies of the US government, nor does it imply that the products identified are necessarily the best available for the purpose.