COA[1] Paper
Secure Data:
Enterprise Information Protection & Control

Introduction

Enterprise Information Protection & Control (EIP&C) addresses:

- The digital management of rights to access information / data
- Control over that information/data - CSMU (copy, store, move and use) of works
- Integrity of the information/data in question.

Information protection needs to cover all data, from word processing documents to data within databases and executable code. It covers enterprise and personal data and is not just confined to protecting entertainment media, which tends to be generically referred to as "Digital Rights Management" (DRM).

Information Control is concerned with the business processes that - for confidentiality, segregation of duties, legal and commercial purposes - tracks rights, rights holders, licenses, sales, agents' royalties and/or associated terms and conditions, using digital technology to apply and enforce the control.

EIP&C does not mandate encryption. Within a secure system, data may be unencrypted and other technologies such as hash functions or watermarking may be adequate if tamper protection or proof of ownership rather than confidentiality is required.

The Problem

In a de-perimeterised world it is generally easier to provide granular levels of data protection. The closer the protection mechanism is to the data, the more effective the protection can be.

In the Jericho Forum design principles[2], commandment #9 states that "Access to data should be controlled by security attributes of the data itself", while commandment #11 states that "By default, data must be appropriately secured when stored, in transit and in use".

On systems under your control, the storage of insecure (typically unencrypted) data that is reliant on system or even network security controls is flawed. The lost PC with client information, or the database administrator who has access to all personal information in a database, are both examples of where the data is inappropriately protected.

For data held outside of your immediate control, there is a need to manage, change or revoke access for data, as well as the need to manage versioning and reduce concurrency.

[1] The Collaboration Oriented Architectures (COA) paper, and associated COA Framework paper and available online at http://www.opengroup.org/jericho/publications.htm
[2] Jericho Forum design principles, #1to #11 – see http://www.opengroup.org/jericho/commandments_v1.2.pdf

Enterprise Information Protection & Control, correctly applied, offers the ability to provide protection and management at the data layer, irrespective of the location of the data.

Current EIP&C solutions are proprietary, limiting their applications by enterprise domain, operating system family, or to specific applications.

Why Should I Care?

The security of the data must reside with that data (Jericho Forum commandment #9) if it is to be adequately protected.

Data leakage arising from personal lapses, process imprecision, network file-shares, FTP (file transfer), E-mail, USB Disk, CD burner, old-CD's, or even floppy disks and output devices (such as printers), is a major issue. Locking down the hardware generally inhibits business, is hard to manage, and is not a viable solution for many organisations.

Today's business depends on an agile and responsive ecosystem, and organisational data regularly needs to exist outside of the organization - with partners, vendors and suppliers, or potentially anywhere that a business relationship exits. Once outside the organisation, it is almost impossible to control, manage access to, or even withdraw access from, that data.

Events change the rights to and classification of data. The need for people to have access to data changes over time. Examples are:

- financial results that go from "top secret" to "in the public domain" overnight
- pre- and post- "embargoed until" times
- price lists that expire at the end of a month
- the person who works for your organisation one day and a competitor the next

Access also needs to change at times of disasters or public emergencies.

EIP&C is a valuable tool for maintaining control of an enterprise's intellectual property while allowing it to be distributed and used across the enterprise's value chain. It's unlikely that all members of the value chain (customers, suppliers, partners, etc.) will have the same IT capabilities or environments, or that an enterprise can dictate those environments. Without neutral standards for EIP&C, this critical technology will fail to aid the extended enterprise in managing its information.

Data that is being properly managed will both control access to that data and define/limit what the person/system is able to do with that data.

Recommended Solution

EIP&C has to be simple to apply and manage in a complex corporate environment (Jericho Forum commandment #2) while solving access to that data, the rights to what you can do with the data, and prevent data leakage - whether accidental, through incompetence, or a malicious act.

It must not rely on a pervasive, ubiquitous real-time network connection (unless this attribute is defined for a particular document), thus enabling off-line working in aeroplanes, remote places in the world, or any other environment where real-time connectivity is not possible.

EIP&C of data is not just about data in discrete files. A data record in a database should also be subject to EIP&C, as could individual e-mails. Also EIP&C should potentially extend to executable files. In a database for example; a personnel file in an HR database should be accessible only by the person it refers to, their line management, and the HR manager for that department, thus ensuring that segregation of duties are adhered to as well as protecting the

privacy of the individual in question. While the database administrator may see, manage, backup, and even port, the data from one database manufacturer to another, they should not be able to read or manipulate the actual data itself.

Background & Rationale

Key escrow and key management

The encryption of documents with a password inevitably leads to documents that cannot be read. The use of EIP&C must enable the management of keys, key escrow and/or access centrally such that access and functionality can be added/changed/revoked simply and easily.

Key management must operate (if allowed by the document EIP&C attributes) in an offline mode.

User identity and the management of users outside of your domain

As EIP&C achieves wide acceptance then it will become increasingly difficult to manage the number of users and devices. Inside the organisation (inside your locus of control) then any EIP&C must interoperate with the organisation's existing authentication system (AD, Kerberos, LDAP, etc.) but also needs to operate effectively with 3rd parties (outside of your locus of control) without compounding the problem, supporting full lifecycle management of those users as part of your EIP&C system. Thus any system must support working in a federated model (or similar) scheme (Jericho Forum commandment #8).

The ease and ability to manage protected documents, potentially coming from multiple organisations, necessitates that client software must be standards-based and capable of interoperating with documents from multiple vendors.

Security functionality

Dependant on the method of accessing/viewing the data, then the security of the endpoint - and thus the ability to trust the operating environment, the user, and where it is actually connected - must be a factor allowing access to the data.

In a trusted computing world, the client program must be capable of being trusted by the data it is processing, to ensure that the data is operating in a valid (not spoofed) environment, such that the client is guaranteed to be able to enforce the attributes required - such as no-copy, no-print, no-screen capture, etc.

Temporal Data classification

Data should be classified, typically by the data originator. However data also has a temporal aspect (commandment #9). Plans for a new product, or stock market results, all change in classification over time. When classifying data it must be possible to specify those temporal conditions.

Auditing of Digital Rights information

Good EIP&C enables audit. Thus EIP&C data and access to that data should be capable of being audited. This is especially important when data is being accessed by systems that are outside of the rights managers control, such as 3rd party systems, or system that are off-line when interoperating with that data.

The linkage of any EIP&C policy manager to the (corporate) directory should ensure adequate segregation of duties on sensitive data (Jericho Forum commandment #10).

Control of data

If EIP&C is to deliver a viable corporate system, then the data "in-the-wild" must be controllable, with the ability to effectively destroy that data (void all access), add and/or change and/or extend access and change the EIP&C attributes of the data.

Such controls must support and integrate into the data information management lifecycle, including support for archiving and the retrieval of EIP&C Data from that archive.

Challenges to Industry

The Jericho Forum believes that there is a lack of agreed standards in this area or that the existing standards are inadequate. If EIP&C is to succeed then 3 things must become standard.

1. The client interface / software must be an open standard

EIP&C must be pervasively available across a wide range of data formats and all platforms. Thus open standards will allow developers to write for and support the widest range of platforms. Open standards ensure that the security principles can be thoroughly reviewed.

Just as it is undesirable / unlikely that any corporation can mandate that another company must install and manage their preferred EIP&C solution, so the prospect of having a client device that has to be installed and support many different EIP&C clients, each with the prospect of interfering with each other or the system, will severely limit widespread adoption.

Instead a single EIP&C client to an open standard will allow wide acceptance (Jericho Forum commandment #2).

2. We must have a standard set of agreed EIP&C classifications

Irrespective of whether EIP&C is applied at document generation time or later in the process, a document must contain (or have access to) all the relevant information / metadata required to process that document. This may be immediate, at some stage in processing or storage, or may be upon leaving the corporation.

To ensure that data can have EIP&C applied then the programs that generate the original data files must be capable of imbedding an agreed standard set of EIP&C Metadata without the need to know the EIP&C product.[3] This will ensure that EIP&C can be applied at a later stage, for example if a document is transferred outside of a secure area/system, or automatically processed when a document is sent to an external e-mail address.

Programs must also be capable of being forced (probably only in a corporate environment) to input EIP&C Metadata (for example, by a flag in a configuration file, or a setting in a MS-Windows registry), ensuring that the entering of EIP&C Metadata can be mandated.

3. We must have a standard for handling In-Clear Classification information

Interoperability presents challenges. Documents under EIP&C control must have enough classification information in-clear to ensure that non-EIP&C systems (such as other programs, storage systems etc.) understand how to correctly handle that document. Such classification information must be protected to ensure that tampering with that in-clear information will be detectable.

---

[3] Dublin Core Metadata Initiative (DCMI) See: http://www.dublincore.org/documents/dcmi-terms/ - ICE (Information and Content Exchange) and the IMS (Instructional Management Specification)
See also http://www.getty.edu/research/institute/standards/intrometadata/

The Way Forward

The continued development of interoperable online metadata standards that support a broad range of purposes and business models enables improved EIP&C.

Existing EIP&C vendors need to collaborate to define a single standard before it's too late for the industry to avoid fragmenting into proprietary point solutions. Fragmented standards lead to fragmented industry and point solutions from which it will be very difficult to recover to open environments.

More open standards need to be defined for EIP&C metadata, the standards and clients used in the management and delivery of EIP&C data.

In particular, there needs to be:

- An open interface / API that can be used to manipulate and query the rights associated with the EIP&C protected data
- A standard, extensible set of classifications as described above
- An open, inherently secure protocol for communicating between consumers of EIP&C protected data and the server or enterprise that controls the data's EIP&C attributes.

Failure to do this will constrain EIP&C technology to a niche market suitable only for internal corporate use. This will destroy its value for leveraging the enterprise-to-enterprise relationships inherent in our de-perimeterized environments.