Il DLP, uno strumento di difesa del patrimonio informativo aziendale

di Michelangelo Uberti

La prevenzione innanzitutto

Tutte le aziende, da quelle più piccole alle grandi multinazionali, hanno oramai compreso l'importanza strategica della difesa del perimetro aziendale, sia dal punto di vista fisico che logico. Anche l'IT Manager meno attento dispone dei più disparati strumenti di difesa attiva e passiva:

- Firewall esterni e interni
- Intrusion Detection System (IDS)
- Intrusion Prevention System (IPS)
- Network Access Control (NAC)
- Antivirus/Antispam per il filtraggio dei messaggi in ingresso
- *etc.*

Appare quindi evidente che l'attenzione è focalizzata su ciò che potrebbe entrare nella rete aziendale mentre invece non viene prestata la dovuta attenzione a ciò che potrebbe uscire.
E' come se gli addetti di un supermercato perquisissero i Clienti all'ingresso ma ignorassero del tutto quelli che escono con le tasche piene di merce rubata!

Per ciò che riguarda il traffico outbound, ogni azienda dispone di strumenti di *web filtering* più o meno evoluti che tipicamente limitano l'utilizzo di determinate porte o protocolli e, in base alle policy aziendali, impediscono l'accesso a siti ritenuti insicuri o non di interesse aziendale. L'utilizzo di un *proxy* per la navigazione rappresenta la misura minima e indispensabile per il controllo del traffico in uscita ma in nessun caso può essere considerata la panacea di tutti i mali.

E' importante ricordare che nessuno strumento di difesa può garantire una sicurezza del 100%, ma con la combinazione di più contromisure è comunque possibile ridurre drasticamente gli incidenti di sicurezza e, qualora questi si verificassero, limitarne i danni.

I dati devono essere considerati al pari degli altri asset aziendali e pertanto devono essere protetti con la stessa tenacia.

Quali dati proteggere?

Ogni azienda dispone, a volte inconsapevolmente, di una vastissima quantità di informazioni che necessita di un elevato livello di attenzione.
Stabilire il tipo di protezione da applicare ad una specifica classe di dati basandosi esclusivamente sul danno che causerebbe la sua compromissione, non è sempre la scelta più giusta: ogni informazione - opportunamente contestualizzata e strumentalizzata - può portare alla perdita di profitto, reputazione e solidità dell'azienda se non addirittura favorire l'insorgere di problemi di natura legale.

Pertanto quali sono i dati che è necessario proteggere? La risposta è una sola: tutti indistintamente.

- Account e dati di profilazione degli utenti interni ed esterni
- Schede del personale e buste paga
- Anagrafica Clienti
- Dati di fatturazione
- Log di ogni tipo (accesso, applicativi, traffico, etc.)
- Repository progetti
- *etc.*

I canali di diffusione delle informazioni

Proteggere un dato non comporta esclusivamente limitarne l'accesso da parte di determinati utenti, ma significa principalmente fare in modo che gli utenti autorizzati rispettino le regole previste dal proprio ruolo. Ciò si traduce, per esempio, nel divieto assoluto di comunicare informazioni riservate ad utenti che non dispongono degli stessi privilegi (es. un funzionario di Human Resources non deve rivelare la retribuzione dei dirigenti ai colleghi, un operatore del Customer Care non deve inviare l'anagrafica Clienti ad un concorrente, etc.). Ovviamente per ottenere questo livello di controllo è necessario utilizzare un adeguato sistema di Identity & Access Management.

Le informazioni possono uscire dal controllo dell'azienda in modi differenti:

- E-mail inviate dalla mailbox aziendale o personale
- Supporti di memorizzazione portatili (pendrive, hard disk esterni, memory card, etc.)
- Perdita/furto degli strumenti aziendali (notebook, palmari, smartphone, etc.)[1]
- Invio di file via FTP o programmi P2P
- Pubblicazione di dati riservati su forum, siti personali, social network, etc.

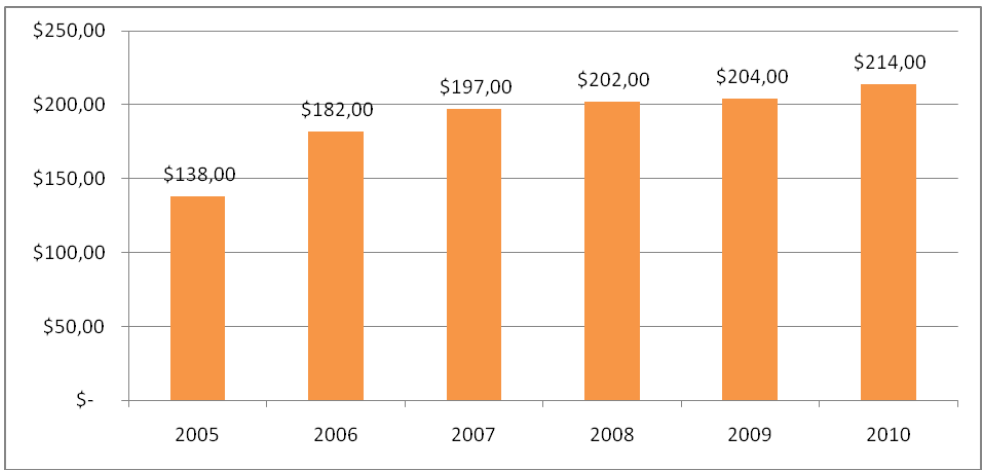
[1] Per la gestione dei dispositivi portatili è altamente consigliato l'utilizzo di suite dedicate al **Device Mobility Management** (DMM). Tali soluzioni consentono ad esempio di effettuare il *wipe* remoto della memoria del dispositivo perso/rubato prima che i dati in esso contenuti vengano compromessi.

- Stampa di documenti classificati

Qualunque sia la modalità con la quale il dato è stato diffuso, la causa afferisce sempre a due categorie: **dolo** e **imperizia**.
Nel primo caso un collaboratore dell'azienda agisce consapevolmente con l'obiettivo di nuocere al business aziendale, nel secondo caso invece l'azione viene effettuata senza comprenderne appieno le conseguenze.
E' doveroso specificare che l'imperizia - ad oggi la quarta tra le cause degli incidenti di sicurezza in azienda - può essere causata sia dall'inadempimento delle procedure previste dalle normative aziendali che dall'effettiva inadeguatezza di tali normative. Per tale motivo è necessario redigere, distribuire e soprattutto applicare delle policy interne più o meno restrittive che aiutino i dipendenti ed i collaboratori esterni a non commettere errori grossolani dovuti all'ignoranza o più semplicemente alla "buona fede".

Pur di risparmiare non si bada a spese!

Risparmiare sulla protezione dei dati aziendali non differisce dal risparmiare sull'installazione di un buon prodotto antivirus o di un sistema antifurto, prima o poi se ne pagano le conseguenze.
Secondo un recente studio del Ponemon Institute[2] il costo per singolo record compromesso negli USA è in costante crescita e nel 2010 ha raggiunto la quota di $214.



Tali stime rappresentano un valido spunto di riflessione in quanto evidenziano il fatto che in caso di una corposa fuoriuscita di informazioni il danno risultante potrebbe essere devastante.
Ad esempio la compromissione del database Clienti (dati sensibili, numeri di carte di credito, dati di fatturazione, etc.) ospitante almeno 100 mila record potrebbe causare un danno superiore ai 21 milioni di dollari!

[2] 2010 Annual Study: U.S. Cost of a Data Breach

Le suddette stime tengono conto delle spese legate alla perdita di business (63%), alla gestione delle controversie legali (24%), alla notifica dell'evento agli utenti coinvolti (7%) e infine all'analisi dell'incidente e alla ricerca (spesso vana) del/i responsabile/i (7%).
Questa è l'ennesima dimostrazione che il detto "prevenire è meglio che curare" è sempre attuale ed è applicabile a qualunque campo di attività.

Le contromisure: il DLP in azione

Il termine DLP, acronimo usato indifferentemente per indicare **Data Loss Prevention** o **Data Leak Prevention**, definisce l'insieme delle tecnologie per l'analisi, l'identificazione, il monitoraggio e la protezione dei dati confidenziali.
La classificazione dei dati avviene sull'analisi dei potenziali "vettori di diffusione":

- **Dati in movimento** (Data in Motion), che includono tutte le informazioni transitanti dalla rete interna verso Internet, ad esempio via e-mail o conversazioni di instant messaging.
- **Dati stanziali ("a riposo")** (Data at Rest), che racchiudono tutte le informazioni contenute in un file system, in un database o in qualsiasi altro strumento di memorizzazione aziendale.
- **Dati in uso** (Data at the Endpoint), che includono tutti i dati memorizzati sui dispositivi remoti (portatili, smartphone, pendrive, etc.).

Data in Motion

Data at Rest

Data at the Endpoint

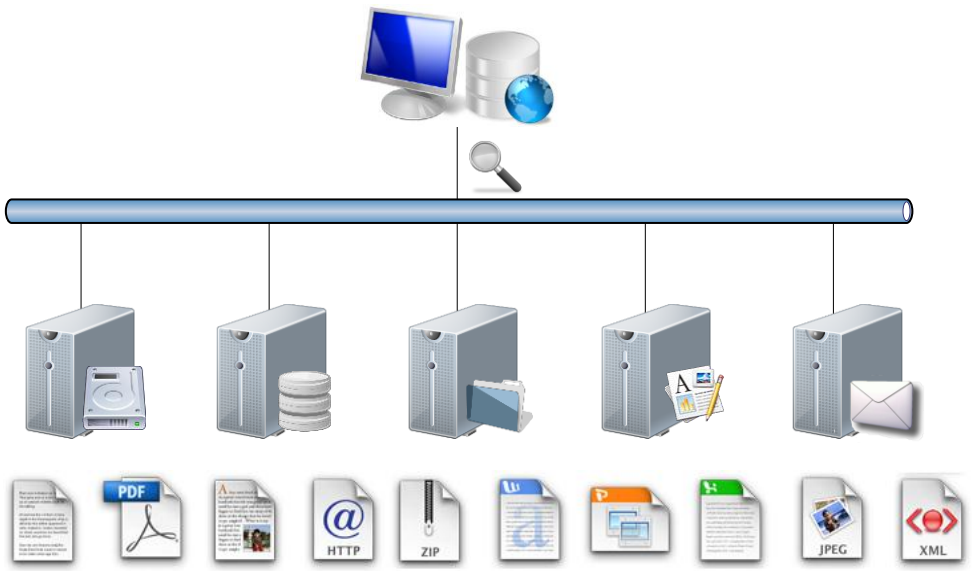
L'obiettivo ultimo del DLP è concettualmente semplice: identificare le sorgenti delle informazioni sensibili, monitorarne tutte le movimentazioni e le eventuali trasformazioni (es. conversioni di formato, estrazione di una quota parte, etc.) ed infine definire quali politiche applicare nelle diverse situazioni verificabili.
Detto così può apparire banale ma in realtà ognuna di queste azioni richiede un'analisi puntuale degli asset aziendali ed una definizione di dettaglio delle politiche da applicare alle singole informazioni.

Le suite dei diversi vendor includono funzionalità simili ma usano spesso terminologie diverse che concorrono a generare confusione nell'utente finale. Nonostante ciò le componenti chiave delle varie suite sono comuni ed includono tre principali funzionalità:

- Data Inventory
- Policy Definition
- Content Monitoring, Filtering & Encryption

Data Inventory – L'identificazione e la classificazione delle sorgenti

Il primo passo da effettuare nel setup di una soluzione di DLP consiste nell'individuazione di tutte le fonti presso le quali sono memorizzati dati di interesse aziendale e quali sono i modi per accedervi. Tale operazione può avvenire mediante una scansione remota dei sistemi "a rischio" – cioè quelli che contengono o potrebbero contenere dati sensibili – oppure mediante l'installazione di un *agent* a bordo delle singole macchine (desktop, laptop, server).



I sistemi interfacciabili sono numerosi ed includono ad esempio:

- Dischi locali dei server
- Database - acceduti via ODBC, JDBC, etc.
- Share di rete - accedute via protocollo CIFS, NFS, etc.
- Applicazioni (es. CMS, CRM, etc.) - fruibili via http
- Mailbox store – accedendo direttamente ai volumi SAN
- *etc.*

Ogni suite di DLP include un modulo delegato al calcolo ed alla memorizzazione del *fingerprint* di un documento, cioè la sua rappresentazione matematica non reversibile. Il fingerprint o *hash* viene tipicamente calcolato sulla base delle informazioni chiave di ogni singolo documento:

- Nome del file
- Dimensione
- Data/ora di creazione ed ultima modifica
- Metadati (autore, titolo, etc.)
- Contenuti

Tali informazioni vengono inserite in un database interno contenente la classificazione delle informazioni rilevate e la catalogazione dei termini chiave in essi contenuti. Il processo di classificazione consente inoltre di definire i pattern per il riconoscimento di singole parti di testo eventualmente estrapolate.

Policy Definition - definizione delle regole di gestione

Il secondo passo prevede la definizione delle regole per la gestione dei dati mediante l'impostazione delle azioni da intraprendere a seguito del rilevamento di una specifica attività. Prendendo ad esempio l'elaborazione di un documento testuale (doc, pdf, etc.) è possibile definire quali azioni effettuare per le principali classi di attività:

Attività	Azione
Stampa su stampante condivisa	Permetti/Blocca/Notifica/…
Copia su periferica USB	Permetti/Blocca/Notifica/…
Invio via e-mail mediante mailbox aziendale	Permetti/Blocca/Notifica/…
Invio via e-mail mediante mailbox personale	Permetti/Blocca/Notifica/…
Print-screen	Permetti/Blocca/Notifica/…
Invio via Instant Messenger	Permetti/Blocca/Notifica/…
…	Permetti/Blocca/Notifica/…

Il livello di granularità nella definizione delle regole è molto elevato e consente ad esempio di:

- Permettere il transito delle informazioni o in generale il completamento dell'azione (es. copia su pendrive);
- Rimuovere i contenuti confidenziali rilevati durante l'invio di un messaggio;
- Bloccare l'invio senza inviare notifiche successive;
- Inserire il dato in quarantena e notificare l'alert agli amministratori;
- Criptare le informazioni ed eventualmente applicare dei DRM per limitarne la lettura o l'elaborazione da parte di utenti non autorizzati.

Tutte le operazioni di definizione delle regole e gestione delle notifiche avvengono tipicamente da una console di amministrazione accessibile via web.

Content Monitoring & Filtering - monitoraggio e applicazione delle regole

L'ultimo passo prevede l'installazione e la configurazione delle sonde dedicate al monitoraggio delle sorgenti e del perimetro della rete.

La prima tipologia analizza e traccia tutti gli accessi effettuati sulle fonti dati in modo da identificare e valutare l'origine delle connessioni e determinare se rispettano i vincoli precedentemente definiti.

La seconda tipologia, tipicamente installata a livello endpoint – cioè sui terminali remoti –, controlla le operazioni effettuate a seguito dell'acquisizione del dato (stampa, modifica, conversione di formato, copia su CD/pendrive, invio via client di posta o webmail, upload via FTP, etc.).

Entrambe le sonde operano in sinergia con i moduli di amministrazione e sono interamente dedicate all'applicazione delle policy definite dagli amministratori IT.

La presa di coscienza

Dopo aver compreso le conseguenze di un potenziale incidente di sicurezza e le diverse contromisure ci sono solo due cose da fare: ***assessment*** e ***scouting***.

L'assessment è un'analisi critica e del tutto onesta sullo stato della propria azienda. Le domande alle quali dare una risposta sono numerose:

- Esistono delle procedure che definiscono il ciclo di vita delle utenze aziendali? Tali procedure vengono applicate a tutti i livelli e indistintamente su dipendenti e terze parti?
- L'azienda dispone di strumenti per l'Identity & Access Management? Tutti i sistemi che gestiscono dati sensibili sono correttamente connessi a tali strumenti?
- La profilazione degli account rispetta il requisito base del "need to know" secondo cui ogni utente deve essere abilitato a visualizzare e gestire solo le informazioni di sua competenza?
- I dati sensibili vengono "securizzati" mediante l'ausilio di strumenti per la cifratura o mediante applicazione di DRM?

- Sono stati messi in campo degli strumenti per la segnalazione e la gestione degli incidenti di sicurezza (ad es. il furto di un notebook)? I collaboratori sono stati informati sulle procedure da seguire in caso di incidenti?
- *etc.*

Le poche aziende italiane che hanno concluso con esito positivo il percorso di certificazione ISO:27001 si troveranno certamente avvantaggiate. Molte altre società dovranno invece avviare una seria riorganizzazione dei propri processi interni.

Lo scouting serve invece ad analizzare quanto offerto dai vendor di tecnologie di sicurezza e valutare quale possa essere la soluzione più adatto alle proprie necessità (ed ovviamente all'ambiente di produzione in cui dovrà operare).
Il miglior consiglio non può che essere uno solo: mettere in campo un Proof-of-Concept e testare accuratamente il prodotto. E' fondamentale valutare se i moduli di inventory sono compatibili con i propri repository, prendere confidenza con le interfacce di gestione, verificare il livello di personalizzazione delle policy di gestione, simulare fughe di dati a livello di endpoint, etc.

Quando si parla di sicurezza il peggior errore che si può commettere è attendere che si verifichi un incidente prima di correre ai ripari. La colpa è certamente maggiore se si era già a conoscenza delle falle nel sistema e si è deciso di non intervenire per risparmiare o perché si è ritenuto che le eventuali conseguenze non fossero poi tanto gravi. Questo concetto è universale ed applicabile non solo all'IT, ma a qualunque campo d'attività.

BABEL S.r.l. - P.zza S.Benedetto da Norcia 33 - 00040, Pomezia (RM) - Tel:. +39 06.9826.9600 - Fax. +39 06.9826.9680
E-Mail: info@babel.it – PEC: babel@pec.babel.it – WEB: http://www.babel.it
Reg. Imprese di Roma N° 06062681009 - N.R.E.A. 953123 - P.I. e C.F. 06062681009 Cap. Soc. € 102.774,00 i.v.
Società soggetta alla direzione e coordinamento della Par-tec S.p.A - Società iscritta al registro delle Imprese di Milano al numero 1293820015