

VINCENZO GAMBETTA

Laureato in Fisica Nucleare, ha trascorso oltre trent'anni in IBM Italia ricoprendo, in Italia e all'estero, incarichi manageriali e professionali. Attualmente è analista di mercato e consulente per la gestione dati, informazioni e documenti presso aziende (fornitrici ed utenti) di medie e grandi dimensioni.

Tra i clienti è Assinform per la quale tiene i rapporti con AIPA/CNIPA, il Ministero dell'economia e delle finanze e le altre istituzioni coinvolte per garantire la miglior applicabilità delle norme sulla conservazione, sulla fatturazione elettronica e sul documento informatico e firma digitale. In Assinform è, anche, coordinatore dei gruppi di lavoro gestione documentale e outsourcing nonché del centro di competenza per le PMI. È autore delle prime due ricerche sul mercato della gestione elettronica dei documenti in Italia, realizzate per ITER, in collaborazione con le Università Cattoliche di Piacenza (1998) e Milano (2000). È abitualmente relatore in seminari e conferenze e scrive per testate specializzate su temi quali gestione elettronica documenti/contenuti, BPM&A, conservazione, fatturazione elettronica, documento informatico e firma digitale. È direttore contenuti del periodico iged.it, specializzato in gestione elettronica dei documenti, e ha la responsabilità scientifica della manifestazione OMAT.

VINCENZO GAMBETTA

LA CONSERVAZIONE DELLA

LA CONSERVAZIONE DELLA MEMORIA DIGITALE

3

MEMORIA DIGITALE



La Fondazione Siav Academy è il Centro per la Ricerca e lo Sviluppo dell'Innovazione Tecnologica nell'ambito di Siav Group.

Attraverso il proprio Comitato Scientifico, costituito da alcuni dei più autorevoli esperti a livello internazionale, la Fondazione intende contribuire al dibattito e all'evoluzione scientifica, normativa e applicativa sui temi della dematerializzazione. La Fondazione, attiva nell'organizzazione di convegni e seminari formativi, opera anche nell'ambito dei Beni Culturali attraverso sponsorizzazioni e progetti dedicati alla conservazione e valorizzazione del patrimonio storico, culturale e artistico.

La Collana "Tecnologia dei processi documentali" rientra nelle attività pubblicistiche che la Fondazione promuove. Curata e coordinata dal Comitato Scientifico, con la supervisione puntuale del prof. Pierluigi Ridolfi, la Collana ha l'obiettivo di approfondire le più importanti e attuali tematiche che riguardano la gestione documentale, per quanto riguarda la normativa, le tecnologie, gli standard e le applicazioni.

in collaborazione con



3



Collana di Minigrafie

Tecnologia dei Processi Documentali

FONDAZIONE SIIV ACADEMY

FONDAZIONE SIAV ACADEMY

VINCENZO GAMBETTA

LA CONSERVAZIONE DELLA MEMORIA DIGITALE

Collana di Minigrafie sulla

Tecnologia dei processi documentali

A cura di Pierluigi Ridolfi

© 2009 Fondazione Siav Academy - Edizione fuori commercio

È consentita la copia di qualunque parte di questo testo purché venga citata la fonte.

ISBN 978-88-97222-03-3



Collana di Minigrafie

Tecnologia dei Processi Documentali

INDICE

Capitolo I *pag. 7*

I concetti base della conservazione

- *Un po' di storia nazionale*
- *Una società fondata sul digitale*
- *I sistemi di conservazione della memoria digitale*

Capitolo II *pag. 25*

I formati e i documenti digitali

- *Cosa si intende per formato?*
- *Tipi di formato*
- *I formati e l'operatività digitale*
- *I formati e la conservazione nel lungo periodo*
- *Alcune ulteriori considerazioni*

Capitolo III *pag. 37*

Il paradosso dell'informazione e del supporto su cui è registrata

Capitolo IV *pag. 43*

I supporti per registrare e conservare le informazioni digitali

- *Tecnologie ottiche: storage professionale e supporti di registrazione consumer, un confronto che dura da 25 anni*
- *Le tecnologie storage disponibili*
 - *Dispositivi a disco*
 - *I Sistemi RAID*
 - *Tipologie di dispositivi a disco*
 - *Le interfacce tra disco e host*
 - *Le tecnologie delle memorie a nastro magnetico digitale*
 - *Lo standard LTO*
 - *Affidabilità e durata nel tempo*
 - *I dischi ottici UDO*
- *Architetture per lo storage in rete*
 - *Infrastrutture di storage scalabili (Grid Storage)*
- *La virtualizzazione*
 - *I livelli di virtualizzazione*
 - *I modelli di virtualizzazione in-band e out-of-band*
 - *Partizionamento*
- *Tiered Storage*
- *Evoluzioni future*

- *Alcuni esempi di realizzazioni RAID*
 - RAID 4
 - RAID 5

Capitolo V

pag. 77

Chi può conservare la memoria digitale?

- *Introduzione*
- *Il problema*
- *Una soluzione*
- *Alcuni suggerimenti*

Capitolo VI

pag. 85

Sistemi di Conservazione Sostitutiva: tecnologia e funzionalità

di Andrea Venturato

- *Gli archivi digitali*
 - *Documenti digitali e archivi digitali*
 - *Gestire la complessità: dallo strumento al sistema*
 - *I rischi degli archivi digitali*
- *Sistemi per la Conservazione Sostitutiva: requisiti e funzionalità di base*
 - *Obblighi del responsabile della conservazione*
 - *Consolidamento del valore probatorio: approfondimento*
 - *Gestione firma digitale multipla*
 - *Gestione marca temporale*
 - *Gestione dispositivi hardware di firma massiva (HSM)*
 - *Gestione processo di riversamento*
 - *Gestione processo di esibizione*
 - *Virgilio, la soluzione per la Conservazione Sostitutiva a lungo termine*
- *Definizione di uno standard di interoperabilità per la gestione dei volumi di conservazione*

Appendice 1

pag. 107

Supporti di memorizzazione a norma per la conservazione

Appendice 2

pag. 113

Deliberazione CNIPA n. 11 del 19 febbraio 2004

Appendice 3

pag. 125

Conservazione dei documenti originali unici

Capitolo I

I concetti base della conservazione

1. Un po' di storia nazionale

In questo campo, come per la firma digitale, in Italia ci si è mossi da pionieri rispetto al resto d'Europa, trascinati dallo stesso entusiasmo e dalla stessa volontà di comprendere e, se non proprio anticipare, almeno seguire da vicino l'evoluzione tecnologica.

L'esigenza di conservare memoria delle attività amministrative e di business comincia a palesarsi nel nostro Paese e a prendere forma di norma con il comma 15 dell'articolo 2 della legge 24 dicembre 1993, n. 537 (finanziaria 1994) che prevedeva che “[...] gli obblighi di conservazione e di esibizione di documenti, per finalità amministrative e probatorie, previsti dalla legislazione vigente, si intendessero soddisfatti anche se realizzati mediante supporto ottico, purché le procedure realizzate fossero conformi a regole tecniche dettate dall'Autorità per l'informatica nella pubblica amministrazione (AIPA)”¹.

¹ L'articolo 4 del decreto legislativo 12 febbraio 1993 n. 39 *Norme in materia di sistemi informativi automatizzati delle amministrazioni pubbliche, a norma dell'art. 2, comma 1, lettera m) della legge 23 ottobre 1992, n. 421* pubblicato nella G.U. n. 42 il 20 febbraio 1993, istituisce presso la Presidenza del Consiglio dei Ministri l'AIPA (Autorità per l'informatica nella pubblica amministrazione) che, nel luglio 2003, viene trasformata in Centro nazionale per l'informatica nella pubblica amministrazione (CNIPA) in forza dell'articolo 176 del decreto legislativo 30 luglio 2003 n. 196.

Il CNIPA, attualmente, opera presso la Presidenza del Consiglio dei Ministri per l'attuazione delle politiche formulate, a nome del governo, dal Ministro per la pubblica amministrazione e l'innovazione.

La sua missione è di contribuire alla creazione di valore da parte della pubblica amministrazione, fornendo a questa supporto nell'uso innovativo dell'informatica e, più in generale, dell'ICT (Information and Communication Technology).

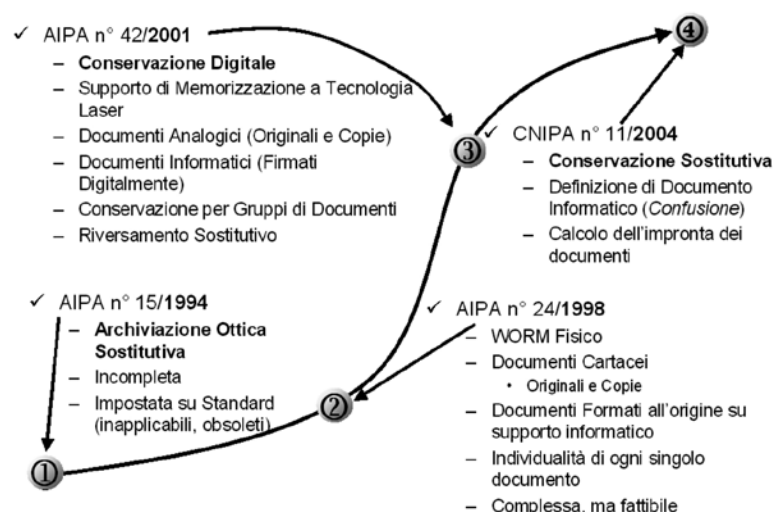


Figura 1 – L'evoluzione delle deliberazioni AIPA/CNIPA sulla conservazione digitale

Inizia così un cammino collettivo di presa di coscienza, di esperienza e di maturazione che porta ai giorni nostri, nei quali il concetto di conservazione della memoria digitale (non solo amministrativa e di business, ma anche storica) anche se non completamente assimilato è, comunque, riconosciuto come una esigenza reale ed imprescindibile.

In questo contesto, le modalità di azione del CNIPA sono:

- l'attività di consulenza e proposta relativamente a strategie e azioni puntuali rivolte ai decisori politici, alle pubbliche amministrazioni e agli operatori del settore;
- l'emissione di normative tecniche a livello secondario, quali linee guida e guide tecniche, sia emesse dal CNIPA, sia da parte delle amministrazioni, con il supporto del CNIPA;
- la valutazione - ex ante (di coerenza con le strategie nazionali di innovazione del Governo), in itinere (durante l'attuazione dei progetti programmati), ed ex post (sui risultati raggiunti) - delle attività ICT della P. A. centrale, sia a livello strategico, sui piani pluriennali di sviluppo dei sistemi informativi, sia a livello di singoli interventi, che per legge sono sottoposti a parere (obbligatorio, ma non vincolante) di congruità tecnica ed economica da parte del CNIPA;
- la definizione e gestione (a livello di committente e con riferimento in particolare alle specifiche funzionali) di progetti dimostrativi, ad alto impatto innovativo, nell'uso dell'ICT nella P.A. centrale, nelle regioni e negli enti locali.

La prima deliberazione AIPA in merito è la n. 15 del 28 luglio 1994².

Questo primo tentativo, pur nella sua incompletezza e nel voler demandare a standard di fatto obsoleti ed inapplicabili le “garanzie” di conservazione nel tempo, ha avuto il merito di essere emanata a pochi mesi di distanza dalla Finanziaria 1994 mantenendo, così, viva l'attenzione sul problema e provocando dialoghi, dibattiti e riflessioni che hanno consentito a tutte le parti in causa di “maturare” in tema di conservazione del digitale.

Si parlava allora di “archiviazione ottica sostitutiva” in quanto il problema al quale si voleva dare una moderna soluzione era la conservazione di copie digitali di documenti cartacei, adottando una tecnica più moderna di quella dei microfilm.

L'impiego dei supporti ottici, che oggi potrebbe far sorridere, era pienamente in linea con le conoscenze e le previsioni delle possibili evoluzioni dell'epoca.

Ha sempre fatto discutere, invece, l'impiego dei CD (ed in seguito dei DVD), ma data la differenza di costo rispetto ai dischi ottici propriamente detti e, visto che si pensava (almeno in questa fase) di dover conservare copie di documenti in formato digitale per tempi dell'ordine di dieci anni, il fatto ha provocato sì diatribe, ma non tali da rappresentare un serio freno.

Bisogna attendere fino al 30 luglio 1998, con la deliberazione AIPA n. 24³, per cominciare a concepire l'attribuzione delle garanzie di inalterabilità nel tempo dei documenti conservati non più ai soli supporti, ma a tecniche quali la firma digitale.

² Pubblicata nella G.U. n 216 del 15 settembre 1994.

³ Pubblicata nella G.U. n 192 del 19 agosto 1998.

L'applicazione di questa deliberazione risultava molto complessa e accusava, da parte degli estensori, una certa "prudenza" dovuta alla carenza di esperienze di applicazione delle tecniche di firma digitale non solo in Italia.

Basti pensare che all'epoca non esisteva ancora una struttura di certificatori⁴.

Questa deliberazione, ancorché come la precedente in pratica poco attuabile, ha avuto i suoi meriti:

- un insieme di definizioni, conseguenti ad un serio ed approfondito esame del problema, tra le quali:
 - la distinzione tra documenti originali e copie,
 - il concetto di *istanza*⁵ andato poi, se non perso, almeno trascurato nelle successive deliberazioni,
 - il concetto di documento formato all'origine su supporto informatico;
- il riconoscimento dell'esigenza di conservare, insieme ai documenti in formato digitale, un elenco di informazioni (metadati) a garanzia della loro originalità ed integrità nel tempo;
- l'individuazione di un responsabile dell'archiviazione.

È la deliberazione 13 dicembre 2001 n. 42⁶ che introduce le norme strutturali di fatto in vigore tutt'oggi.

⁴ I *certificatori* sono terze parti fidate il cui compito principale è quello di associare ad ogni chiave privata, di una coppia di chiavi asimmetriche, un certificato che, utilizzato in congiunzione alla chiave pubblica della coppia, consente di verificare la validità della firma ed attribuirla ad un titolare ben identificato.

L'Elenco pubblico dei certificatori è stato istituito con la circolare AIPA del 13 luglio 2000 n. 26, pubblicata in G.U. il 20 luglio del 2000.

⁵ L'istanza di un documento registrato è il risultato di una operazione di archiviazione effettuata a fronte del corrispondente documento d'origine. Ciascuna istanza di un documento registrato è individuata, nell'ambito di questo, dal numero d'ordine con cui è stata generata.

⁶ Pubblicata nella G.U. n. 296 del 21 dicembre 2001.

Vista l'evoluzione normativa e tecnologica, è introdotta la possibilità di portare in conservazione i documenti informatici⁷ (firmati digitalmente) e, forse per questo motivo, non si parla più di archiviazione ottica sostitutiva bensì di *conservazione digitale*.

Conseguenza della gestione dei documenti firmati digitalmente è l'introduzione del concetto di riversamento sostitutivo, ossia del processo che trasferisce uno o più documenti conservati da un supporto di memorizzazione ad un altro modificandone la rappresentazione digitale.

Le altre principali novità sono rappresentate da:

- le definizioni di documenti analogici (cartacei) originali unici e non unici (definizione questa che ha trovato molti fieri oppositori, divenuta poi effettiva con il Codice dell'amministrazione digitale - decreto legislativo 7 marzo 2005, n. 82⁸);
- l'impiego della tecnica della firma digitale non più su ogni singolo documento, ma su un insieme (gruppo) di documenti;
- l'abbandono del concetto di supporto di memorizzazione a tecnologia laser "con modifica fisica ed irreversibile del supporto in fase di registrazione" (si riconosce la validità delle tecniche di firma digitale per garantire autenticità e inalterabilità nel tempo, mentre al supporto resta il compito di assicurare la massima durata possibile delle registrazioni fisiche);

⁷ Il documento informatico così concepito è stato introdotto nel nostro ordinamento dal comma 2 dell'articolo 15 della legge "Bassanini I" (n. 59 del 15 marzo 1997, pubblicata nella G.U. n. 63 del 17 marzo 1997 - Supplemento ordinario n. 56), mentre ha trovato una regolamentazione compiuta con il DPR 28/12/2000, n. 445, pubblicato in G.U. il 20 febbraio 2001.

⁸ Pubblicato nella G.U. n. 112 del 16 maggio 2005 - Supplemento ordinario n. 93.

- la possibilità di impiegare supporti differenti da quelli ottici.

Si era, dunque, in presenza di un quadro normativo sufficientemente consistente e chiaro che avrebbe potuto consentire la realizzazione di applicazioni pratiche.

Nel frattempo il concetto di firma digitale, e più estesamente di firma elettronica, si stava affermando in tutto il mondo ed in particolare nell'Unione Europea; ciò ha portato alla direttiva 1999/93/CE del Parlamento europeo e del Consiglio del 13 dicembre 1999 relativa ad un quadro comunitario per le firme elettroniche, recepita nel nostro ordinamento con il decreto legislativo n. 10 del 23 gennaio 2002 e successivo DPR 7 aprile 2003, n. 137 - Regolamento recante disposizioni di coordinamento in materia di firme elettroniche a norma dell'articolo 13 del decreto legislativo 23 gennaio 2002, n. 10.

Il documento informatico diventa pertanto molto più genericamente “la rappresentazione informatica di atti, fatti o dati giuridicamente rilevanti”; mentre per quanto riguarda la loro sottoscrizione si passa ad uno spettro di modalità che va dalle semplici firme elettroniche (“l'insieme dei dati in forma elettronica, allegati oppure connessi tramite associazione logica ad altri dati elettronici, utilizzati come metodo di autenticazione informatica”) alle firme elettroniche qualificate (“la firma elettronica ottenuta attraverso una procedura informatica che garantisce la connessione univoca al firmatario e la sua univoca identificazione, creata con mezzi sui quali il firmatario può conservare un controllo esclusivo e collegata ai dati ai quali si riferisce in modo da consentire di rilevare se i dati stessi siano stati successivamente modificati; essa inoltre deve essere accompagnata da un certificato qualificato e creata mediante un dispositivo sicuro per la creazione della firma”); la

firma digitale già prevista nel nostro ordinamento non è altro che uno dei modi in cui si può realizzare una firma elettronica qualificata.

Si arriva così alla deliberazione CNIPA 19 febbraio 2004, n. 11 che, in sostanza, riproduce la deliberazione 13 dicembre 2001 n. 42 adeguandola alla direttiva europea relativa ad un quadro comunitario per le firme elettroniche, oltre a qualche aggiustamento secondario.

Si coglie, altresì, l'occasione per rinominare ancora il processo passando da conservazione digitale a *Conservazione Sostitutiva*.

Quest'ultimo termine risulta alquanto ambiguo⁹, specialmente se si considera che lo si usa anche per la “conservazione” dei documenti informatici.

Ciò non ha, però, impedito che le emanazioni praticamente simultanee della deliberazione 11/2004, del decreto del Ministro dell'economia e delle finanze 23 gennaio 2004¹⁰ e del

⁹ Si noti che la *Conservazione Sostitutiva*, della quale peraltro non si fa cenno nel Codice dell'amministrazione digitale, viene utilizzata solo nella deliberazione CNIPA n. 11 del 2004. In realtà è più corretto intenderla come processo somma di due processi distinti e consecutivi: quello di “riproduzione sostitutiva” e quello di “conservazione”; di per sé, la conservazione sostitutiva, di fatto, non esiste. La *riproduzione sostitutiva*, molto più correttamente, la si può intendere di due tipi, del tutto diversi tra loro:

- il processo con il quale un documento analogico (ad esempio, un documento cartaceo) dà origine a un documento informatico (ad esempio, l'acquisizione di una pagina a stampa mediante scanner, che crea un'immagine in formato quali .jpg o .tif);
- il processo con il quale un documento informatico dà origine a un altro documento informatico con modifica del formato.

In entrambi i casi, pur essendo cambiato il modo in cui il documento viene rappresentato, il suo contenuto informativo resta nella sostanza inalterato.

Anche se si tratta di una terminologia molto diffusa nel linguaggio corrente, occorrerà evitarne l'utilizzo in quanto si tratta di un concetto impreciso.

¹⁰ “Modalità di assolvimento degli obblighi fiscali relativi ai documenti informatici ed alla loro riproduzione in diversi tipi di supporto”.

decreto legislativo 20 febbraio 2004, n. 52¹¹, con le relative circolari esplicative dell'Agenzia delle entrate, muovessero finalmente il mercato verso un numero sempre più significativo di realizzazioni di applicazioni di Conservazione Sostitutiva.

Si possono, in conclusione di questa breve storia, fare alcune considerazioni:

- il processo di conservazione di informazioni in formato digitale è oggi un concetto acquisito ed è ritenuto quantomeno utile, se non indispensabile, da aziende, amministrazioni ed organizzazioni;
- il processo in questione è estremamente complesso e le componenti tecniche, logiche e procedurali in gioco sono tali e tante che non ci si può aspettare di individuarle in poco tempo e averle, in tempi brevi, assimilate da parte di tutte le parti in causa; a questa situazione si somma l'evoluzione tecnologica che mostra dinamiche fino ad oggi sconosciute;
- è pertanto importante procedere nella realizzazione di applicazioni di conservazione della memoria digitale al meglio delle possibilità tecnologiche, organizzative ed economiche; ciò vale in particolare per le organizzazioni di minori dimensioni in quanto le grandi già si sono attrezzate in merito. In effetti è quello che sta oggi accadendo;
- non deve sembrare poco il punto in cui siamo (conservare insieme di documenti in lotti/gruppi le cui dimensioni sono legate al supporto utilizzato o, comunque, alle cadenze scelte). Basta riflettere un attimo sulla nostra strut-

tura produttiva¹² e su come sia oggi diffuso e recepito dalla maggior parte delle organizzazioni il concetto di Conservazione Sostitutiva.

I tempi sono ora maturi per fare ulteriori passi in avanti ed emanare regole tecniche più moderne che rimpiazzino la deliberazione CNIPA 11/2004 che, come si è visto, ha molto contribuito a far decollare nel nostro Paese i processi di conservazione della memoria digitale, ma che risulta ormai non più adeguata ai tempi, alle tecnologie ed alle esigenze.

2. Una società fondata sul digitale

La società moderna sta sperimentando un'esplosione di informazioni in formato digitale, ben superiore anche a quella demografica ipotizzata per i prossimi cinquant'anni: entro il 2050 la popolazione mondiale si prevede passi dagli attuali 6,5 miliardi a ben 9 miliardi di persone.

Creare e accedere ai dati in formato digitale non è mai stato peraltro così facile.

Oltre un miliardo di utenti crea dati su Internet; questi crescono di giorno in giorno, trasformando la Rete in un immenso archivio di contenuti digitali.

Un altro fattore che ha contribuito al rapido accumulo di dati è il proliferare di dispositivi mobili (telefoni cellulari, computer portatili e palmari) che consentono di acquisire e trasferire, mol-

¹¹ "Attuazione della direttiva 2001/115/CE che semplifica ed armonizza le modalità di fatturazione in materia di IVA".

¹² Nel contesto italiano le aziende di minori dimensioni (da 1 a 49 addetti) rappresentano il 99,3% del totale. Tra queste, le imprese più piccole, ossia quelle con numero di addetti inferiori a 20 unità, sono oltre il 98% dell'intera categoria (oltre 3.850.000 aziende). Un andamento non molto dissimile lo si ritrova analizzando le pubbliche amministrazioni.

[Elaborazioni 2003 Assinform/NetConsulting su dati ISTAT 2001]

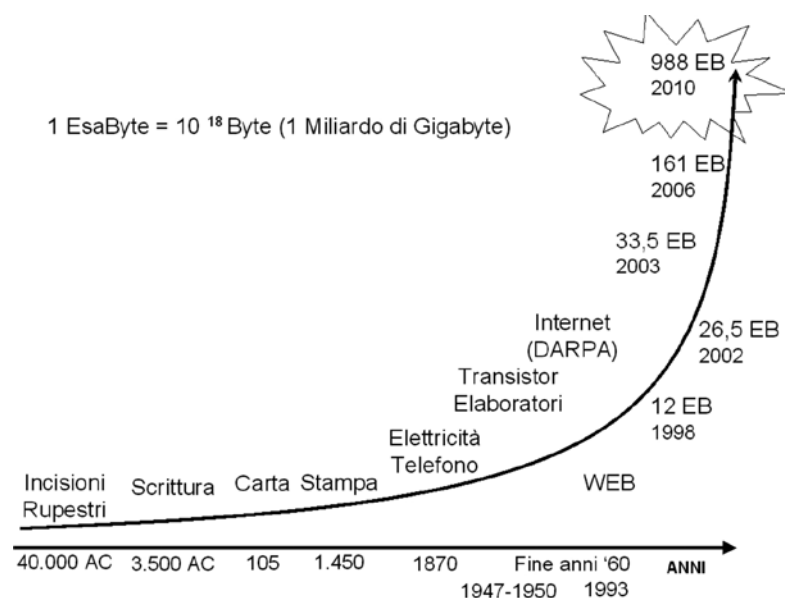


Figura 2 - La crescita delle informazioni
[Università di Berkeley (CA) – (2003) - IDC per EMC Corporation (2007)]

to facilmente, dati tramite reti private, reti wireless, Internet, ... senza contare i dispositivi di raccolta e monitoraggio automatico dei dati che, praticamente sconosciuti 50 anni fa, potrebbero tra non molto superare in numero la popolazione mondiale.

Possiamo datare il momento storico del passaggio dal supporto cartaceo al digitale tra il 1993 ed il 1998; non è casuale che nello stesso periodo i costi di memorizzazione su disco magnetico siano diventati più convenienti che non quelli su carta o microfilm.

Da allora, la produzione e il conseguente impiego di contenuti digitali, ha preso a crescere in modo impressionante: dai 12 esabyte (miliardi di miliardi di byte) prodotti nel 2002, ai 161

Esabyte del 2006, ai 988 che si prevede saranno prodotti nel 2010.

L'ultimo aggiornamento, sempre di IDC, prevede che, nel 2011 le informazioni prodotte saranno ben 1.800 Esabyte (leggasi 1,8 Zettabyte), con un aumento di circa il 5% della crescita delle informazioni prodotte negli ultimi quattro anni di analisi (dal 2007 al 2011 nei confronti della crescita prevista dal 2006 al 2010).

Nello stesso anno si ipotizza che l'85% delle informazioni generate dovranno, per un motivo o per l'altro, essere gestite da aziende, organizzazioni o amministrazioni che avranno l'onere di preoccuparsi di tutto ciò che ne consegue (affidabilità, sicurezza, privacy, conservazione, conformità alle norme vigenti e così via).

Questi contenuti digitali sono i veri motori del moderno business e consentono di ottenere i vantaggi competitivi (in termini sia di efficienza sia di efficacia) indispensabili a stare al passo con i tempi.

Aziende e organizzazioni fondano, dunque, il loro business su informazioni in formato digitale; vi è, poi, la consolidata tendenza (o esigenza) a convertire in digitale sempre maggiori quantità di documenti cartacei e ad impiegare sempre di più documenti che esistono e sono gestiti solo in formato digitale.

L'informazione digitalizzata non solo aumenta la capacità di comunicare, di condividere e di fruire delle informazioni, ma, opportunamente controllata, può essere elaborata a tutto vantaggio dell'economia e della velocità dei processi produttivi, amministrativi e di business (*efficienza*, dunque) consentendo di rendere più veloce ed automatico tutto ciò che non richiede intervento umano.

Un'ultima conseguenza è rappresentata dal potenziale che si genera per la creazione di nuovi prodotti e per la fornitura di nuovi servizi (in poche parole, non solo efficienza, ma anche *efficacia*).

Siamo pertanto in un momento di transizione molto importante che consentirà a chi lo saprà cavalcare di affermarsi nel proprio contesto di riferimento.

La vera sfida di questi anni è e sarà dunque quella di riuscire a gestire e conservare in modo efficiente le informazioni digitali e, conseguentemente, ad impiegarle in modo efficace.

Non è impresa semplice in quanto *la nostra capacità di generare e trasmettere informazioni ha superato di gran lunga la capacità di ricercare, ritrovare e disporre delle stesse*.

Non solo, ma dovremo imparare anche a conservare memoria storica e/o amministrativa della nostra attività quotidiana e, pertanto, imparare a gestire la conservazione della memoria digitale.

È impensabile, come si dava per scontato solo alcuni decenni fa, conservare memoria dei nostri tempi in una forma diversa da quella con cui si opera (ossia digitale).

Il problema della conservazione della memoria digitale s'impone dunque e, come si può facilmente intuire anche da parte dei non addetti ai lavori, richiede un approccio multi-disciplinare che deve adottare più tecnologie, eventualmente integrandole tra di loro.

Una cosa è comunque certa: non si torna indietro!

3. I sistemi di conservazione della memoria digitale

Il problema è certamente molto complesso. In tutto il mondo stanno prendendo corpo, con enorme impiego di risorse, importanti studi e progetti per individuare come riuscire a conservare per decenni, se non per secoli, l'ingente quantità di differenti tipologie di informazioni digitali che abbiamo generato e che genereremo.

Si tratta di sistemi di conservazione digitale (*Digital Preservation Systems*) che hanno l'obiettivo - tramite opportuni processi, attività e tecnologie - di garantire la sicurezza e l'autenticità delle informazioni digitali nonché di rendere possibile l'accesso alle stesse e la loro comprensione in un futuro, anche remoto, consentendo:

- di rileggere correttamente i bit così come erano stati originariamente registrati, su di un supporto molto probabilmente differente da quello attuale (*Bit Preservation*)
- di mantenere la comprensione e le modalità d'impiego dei dati anche quando le attuali tecnologie impiegate per l'hardware, i sistemi operativi, i sistemi di data management e gli applicativi potrebbero non essere più disponibili (*Logical Preservation*).

Si può provare a schematizzare i quattro elementi fondamentali su cui poggiano questi sistemi; essi sono:

- *Il formato*
Alla base della capacità d'interpretare un documento le informazioni conservate è il "formato", con cui sono state portate in conservazione, che deve essere immutabile e, a garanzia della fedeltà di presentazione, deve "contenere" tutte le caratteristiche utili alla completa comprensione del documento.

A fronte dell'imprevedibilità dell'evoluzione delle tecnologie digitali esso deve, inoltre, corrispondere fedelmente ad uno standard interamente descritto e le cui specifiche siano state rese di dominio pubblico.

- *La garanzia dell'integrità e dell'autenticità delle informazioni conservate*

Opportune combinazioni di tecniche di firma digitale e di buone pratiche di controllo e custodia (ad esempio l'ininterrotta custodia) consentono poi di avere la ragionevole certezza che le informazioni che si andranno ad esaminare saranno affidabili. Questo è forse l'aspetto più delicato dal punto di vista normativo che deve riconoscere/attribuire validità giuridico/legale ai documenti conservati secondo modalità e processi e regole che garantiscano che le informazioni conservate mantengano nel tempo l'integrità e l'autenticità che avevano quando il sistema di conservazione le ha prese in carico. Tali regole, peraltro, non devono essere troppo legate alle tecnologie correnti per evitare che la loro evoluzione renda vano il lavoro precedentemente fatto.

- *I metadati*¹³

Col crescere delle tipologie e della quantità delle informazioni da conservare si rafforza, inoltre, (non che sia mai venuta meno!) l'esigenza di definire quali

sono i metadati da conservare per comprendere appieno il contesto cui l'informazione si riferisce, a quale altre informazioni è correlata ed in che modo, ... Essenziali divengono quindi le *regole* - comprensive di quelle di ritenzione - che l'organizzazione/azienda deve definire per gestire l'intero ciclo di vita del documento/informazione. Va da sé l'esigenza di modificare e tracciare i metadati stessi qualora si presentasse l'esigenza di modificarli.

- *L'integrità delle registrazioni*¹⁴

Per quanto riguarda, infine, il mantenimento nel tempo dell'integrità delle registrazioni fisiche sarà necessario imitare il comportamento di tutte quelle aziende che ormai di digitale vivono e non si possono permettere di perdere anche una sola piccola parte del proprio patrimonio informativo, ossia:

- da un lato procedere con organiche modalità di gestione che garantiscano di non perdere informazioni, affidandosi ad un opportuno mix di tecnologie e di operazioni di duplicazione e replica, di refresh, di back-up/recovery e di disaster recovery;
- dall'altro, grazie ad appropriate tecniche di virtualizzazione e a più livelli di storage (*Tiered Storage*), garantirsi (cosa oggi praticamente possibile,

¹³ La più semplice definizione di metadato è "dato relativo al dato"; esso è pertanto un'informazione che descrive, o completa, il dato cui si riferisce.

I metadati, ad esempio, indicano la data in cui un documento è stato generato/archiviato e chi ha eseguito tali operazioni e possono essere utilizzati per una successiva ricerca.

I metadati sono fondamentali nel definire il contesto cui un dato, documento o informazione si riferisce e come questi possono essere utilizzati.

I metadati possono essere memorizzati con il singolo dato o documento cui si riferiscono o separatamente.

I metadati possono dare risposte a quesiti semplici o molto sofisticati.

¹⁴ A proposito dell'integrità delle registrazioni lo scrivente ha redatto una guida per la scelta delle tecnologie dei supporti di memorizzazione digitale per la conservazione e con l'indicazione di una serie di procedure minime da utilizzare per garantire la protezione dei dati nel lungo periodo. Questa guida è stata scritta per conto di Assinform (l'associazione nazionale - aderente al sistema Confindustria - delle principali aziende di Information Technology operanti sul mercato italiano) ed è il risultato della collaborazione di Dell, EMC2, Hewlett Packard, Hitachi Data Systems, IBM e Sun Microsystems. La guida è scaricabile dalla home page del sito di Assinform all'indirizzo: www.assinform.it/download/allegati08/Guide_Tecniche_articolo_11_Regole_tecniche.zip.

indipendentemente dal fornitore) la disponibilità di un “contenitore” dalle dimensioni virtualmente illimitate che non ponga limiti alle quantità di dati da conservare/gestire e consenta di utilizzare contemporaneamente tecnologie e piattaforme differenti, con la possibilità di modificare, cambiare, o stravolgere completamente, l’infrastruttura adottata, mantenendo invariata l’accessibilità alle informazioni.

Senza pretesa di essere esaustivi, sono di seguito elencate alcune delle prerogative, oltre alla corretta generazione e gestione dei metadati, che un sistema di conservazione deve avere al fine di *garantire nel tempo la leggibilità, l’integrità e l’autenticità delle informazioni conservate*:

- appropriate politiche e procedure a garanzia:
 - della sicurezza fisica e logica,
 - della business continuity,
 - del disaster recovery,
 - dell’appropriato controllo degli accessi;
- rigorosi processi e procedure di presa in carico delle informazioni/documenti da conservare in grado, tra l’altro, di svolgere, con la massima attendibilità possibile, tutte le verifiche necessarie su certificati, marche temporali, CRL (*Certificates Revocation List*);
- un ambiente sicuro per l’apposizione di firme digitali;
- auditabilità e tracciabilità delle procedure e dei processi.

Appare evidente la complessità di quanto dovrebbe essere messo in opera per garantire la conservazione della memoria digitale.

Ciò premesso, non significa che si debba attendere la realizzazione di sistemi di conservazione digitale completi ed evoluti

per cominciare ad operare in ottica di protezione delle informazioni (digitali) che caratterizzano la civiltà attuale.

Gli strumenti, anche se non perfetti come li desidereremmo, sono disponibili tanto è vero che, pur nei limiti caratteristici di un momento di transizione, le grandi aziende, organizzazioni ed amministrazioni si sono già attrezzate per garantire la protezione della propria memoria digitale.

Le piccole sono più in difficoltà, se non proprio assolutamente impreparate, e per conservare i loro dati in formato digitale si stanno sempre più affidando a servizi esterni.

In questa sede ci si focalizzerà sui formati e sull’integrità delle registrazioni, lasciando ad altra sede una più approfondita analisi degli altri due punti.

Capitolo II

I formati e i documenti digitali

1. Cosa si intende per formato?

Quando si ha a che fare con informazioni digitali non si può prescindere dal loro “formato”, ossia del modo in cui questo tipo di informazione viene strutturato e registrato sul supporto di memorizzazione.

Il formato è in genere composto da un corpo che contiene le informazioni e dai dati operativi che consentono ad opportuni software di elaborare e presentare all’essere umano, in modo comprensibile e strutturato, le informazioni stesse; al corpo è, in genere, associato un insieme di metadati (testata o header), ovvero le informazioni necessarie ad interpretare correttamente il suo contenuto.

È facile incontrare, poi, situazioni in cui un formato è in grado di contenerne (incapsularne) altri; ad esempio un formato utilizzato per gestire testo che contenga anche immagini o grafici.

È importante riconoscere lo stretto legame esistente tra il software che consente di generare l’informazione, modificarla e manipolarla, di strutturarla e di visualizzarla (o più genericamente di renderla fruibile all’essere umano) ed il formato in cui essa è organizzata e registrata.

Tra le funzioni di tali software (o programmi) va evidenziata quella che serve a visualizzare (rendere fruibile) il contenuto informativo (digitale); questa funzione può essere svolta anche da programmi esplicitamente dedicati a questo scopo.

In genere un programma ha un formato *preferenziale* con il quale salva i documenti/informazioni che è in grado di generare ed un certo numero di formati *compatibili* che, in altri termini, può riconoscere e gestire in input e/o in output.

È facilmente intuibile che questo tipo di “conversione” tra formati (di fatto è ciò che avviene tra i formati compatibili e quello preferenziale), che come altre funzionalità può rilevarsi importante nell’operatività di tutti i giorni, è potenzialmente fonte di perdita di informazione.

Ciò è dovuto al fatto che non sempre si è in grado di operare una conversione completa; col trascorrere del tempo il fenomeno, inevitabilmente, si enfatizza in quanto da un lato le tecnologie si evolvono e dall’altro aumentano le esigenze degli utenti: di conseguenza, ogni programma tende a specializzarsi e a dotare di nuove funzionalità i formati preferenziali e le modalità con cui questi sono gestiti.

Un’altra importante caratteristica di un formato è, poi, il suo livello di *portabilità* che è funzione di due componenti:

- il numero di programmi in grado di gestirlo;
- il numero di piattaforme in grado di garantire il funzionamento dei suddetti programmi.

2. Tipi di formato

L’evolversi delle tecnologie e la crescente disponibilità e complessità dell’informazione digitale induce nell’utente l’esigenza di gestire sempre maggiori forme di informazione digitale (testo, immagini, filmati, formati che ne incapsulano altri, ...) e di disporre di funzionalità sempre più sofisticate e specializzate per renderne più facile la creazione, la modifica e la manipolazione.

Questo fenomeno porta all’aumento del numero dei formati disponibili e dei corrispondenti programmi necessari a gestirli nonché delle piattaforme su cui questi operano.

Altra variabile da considerare è se il formato, in cui l’informazione digitale è gestita e conservata, è proprietario o aderisce ad uno standard, meglio se formalmente riconosciuto: oggi esistono oltre 15.000 estensioni (*file name extension*), che individuano formati in cui sono generate e conservate informazioni digitali, e di queste solo per un migliaio è disponibile una “qualche forma” di descrizione.

A completamento di questa ricognizione, possiamo elencare ed operare una prima sommaria, e non esaustiva, catalogazione dei più diffusi formati (non è questa la sede per approfondire le classificazioni sui formati che peraltro vedono divergenze anche tra i più diffusi e quotati metodi di classificazione):

- immagini (GIF, JPG, BMP, TIF, CGM, EPS, SVG, ...)
- suoni (MP3, WAV, ...)
- filmati (MPG, MPEG, AVI, WMV, ...)
- testi/documenti (DOC, HTML, PDF, ...)
- dati binari, più o meno strutturati, come, ad esempio, gli output di strumentazione di laboratorio
- eseguibili (EXE, ...)
- calcolo (XLS, MDB, ...)
- archiviazione e compressione (ZIP, RAR, ...)

Tra i formati non citati vale la pena di ricordare quelli relativi alle e-mail che meritano riflessioni ed approcci del tutto particolari.

I problemi che inducono la complessità ed il numero di formati e la loro gestione sono dunque di dimensioni tanto vaste che i

molteplici progetti di valenza internazionale non sono ancora riusciti a venirne a capo. Ciò non deve indurre però ad una attesa passiva e rassegnata.

3. I formati e l'operatività digitale

Si può innanzi tutto fare una prima riflessione: l'informazione digitale consente di essere facilmente memorizzata, vi si può altrettanto facilmente accedere e riutilizzarla, modificarla e manipolarla, in altre parole, elaborarla ed ottenere nuova informazione.

Come accennato, a questi fini sono disponibili un numero elevato di formati e di programmi in grado di gestirli.

Questi formati e i programmi che li gestiscono, che sono poi quelli che consentono e facilitano l'operatività giorno per giorno sul digitale, vanno valutati in funzione di alcune caratteristiche, quali:

- la *diffusione*, ossia il numero di persone ed organizzazioni che li adotta;
- la *portabilità*, ancor meglio se essa è indotta dall'impiego fedele di standard documentati e accessibili;
- le *funzionalità* che l'utente ha a disposizione per elaborare l'informazione e collegarla ad altre (ad esempio la gestione di link);
- la capacità, per i programmi, di gestire contemporaneamente un *numero congruo* (in funzione delle esigenze dell'utente) di formati;
- la diffusione di *visualizzatori* che consentono una fruibilità delle informazioni in essi contenute indipendentemente dalla possibilità di rielaborarle.

Altre caratteristiche importanti sono la capacità di occupare il minor spazio possibile in fase di memorizzazione dell'informazione (a questo proposito vanno valutati, in funzione delle esigenze dell'utente, gli eventuali livelli di compressione utilizzabili) e la possibilità di gestire il maggior numero possibile di metadati, compresi i riferimenti a chi ha eseguito modifiche o aggiunte.

È facilmente comprensibile come, nella fase di gestione del digitale, l'utente debba avere a disposizione la massima flessibilità possibile in termini di formati e funzionalità disponibili.

Gli unici limiti sono quelli che un'organizzazione impone a se stessa quando, per esigenze di interscambio ed interoperabilità, può determinare i formati, e i relativi programmi di gestione, che maggiormente soddisfano le contingenti esigenze operative.

4. I formati e la conservazione nel lungo periodo

Quando si tratta di conservazione nel lungo periodo, i principi ispiratori e gli elementi di valutazione sono ben diversi.

Innanzitutto si ha a che fare, normalmente, con documenti di valore storico o con business record ossia "informazioni create, ricevute e conservate come prova o dato sensibile, da una persona od organizzazione, in transazioni di business o nel rispettare obblighi di legge" (standard ISO 15489).

Si tratta, dunque, di informazioni che in qualche momento del loro ciclo di vita sono state definite, o riconosciute, come informazioni che devono essere conservate, così come erano in quel momento, per un periodo superiore ad un certo numero di anni. Tali informazioni devono essere protette da qualunque modifica, sia essa originata da possibilità intrinseche al forma-

to che le accoglie, quali macroistruzioni o codici eseguibili in grado di attivare funzionalità che possano modificare gli atti, i fatti o i dati nello stesso rappresentati, o da interventi esterni di qualunque natura (fraudolenti o meno).

È dunque necessario garantire l'integrità delle informazioni, ossia che esse non vengano alterate e/o distrutte e pertanto siano, *in ogni momento*, corrispondenti a quelle che si è deciso di conservare allorché si è presa tale decisione e che, pertanto, siano anche legalmente veritiere.

Dato per scontato che si sia in grado di mettere in atto tutti quegli accorgimenti e quelle procedure che consentiranno, in un futuro, di rileggere le informazioni interessate così come erano state registrate (opportuna gestione dello storage) e che l'integrità delle informazioni sia sempre verificabile/controllabile (ad esempio con l'uso appropriato delle tecniche di firma digitale e di processi che siano in grado di assicurarne nel tempo l'autenticità e l'integrità delle informazioni) quale è il contributo del formato che accoglie le informazioni ai fini della loro conservazione nel lungo periodo?

Si dà, inoltre, per scontato che siano disponibili, e messe in pratica, tutte le opportune pratiche archivistiche atte a descrivere il contesto del documento, il livello di importanza che esso riveste per l'organizzazione che lo detiene e, se applicabile, quando può (deve) essere eliminato in quanto non più utile.

Date le premesse - molteplicità dei formati, caratteristiche delle informazioni da conservare e complessità di gestione e dinamicità di evoluzione del digitale - *conviene iniziare a ragionare per gradi, senza pensare di prendere in considerazione tutte le possibili casistiche.*

In prima istanza, senza voler sminuire le esigenze di conservazio-

ne di altre tipologie d'informazione e nella convinzione di coprire una parte molto consistente delle situazioni, si ritiene sia bene concentrare l'attenzione sui formati che consentono di presentare le informazioni *sotto forma di documento tradizionale*, ossia di formati con testo, grafici ed immagini che simulano la struttura e le tecniche di rappresentazione di un documento cartaceo.

Sembra quasi superfluo ricordarlo, ma un formato che soddisfi suddette esigenze di conservazione è anche adatto a rappresentare documenti ai quali apporre una firma digitale, vuoi per sottoscrizione vuoi come strumento per garantirne l'autenticità e l'origine.

Il fatto di adottare formati da utilizzare per l'operatività di tutti i giorni o formati, in cui convertire i documenti, che più si adattano a garantire una conservazione il più accurata possibile non deve meravigliare, né preoccupare.

Proprio poiché ci si sta riferendo ad attività legate all'operatività corrente saranno sempre disponibili tecnologie e strumenti che consentiranno una conversione tra formati, senza perdita o alterazione di contenuto informativo.

Questo approccio consente anche di avere sempre a disposizione i formati più consoni e produttivi in funzione delle esigenze operative e delle tecnologie disponibili.

Di seguito sono elencate le caratteristiche che si ritiene debba avere un formato per offrire il proprio contributo alla conservazione nel lungo periodo dell'informazione (memoria) digitale, in altri termini per essere "riconosciuto" e utilizzato correttamente nel tempo:

- *Immodificabilità*: il formato non deve consentire di apportare modifiche al contenuto informativo (a qualunque livello, compresa struttura e modalità di visualizzazione).

- Ne consegue anche l'esigenza di poter sempre disporre dell'intero contenuto informativo originale; ciò porta a non preferire formati immagine che prevedano compressione con perdita di informazione (vedansi di seguito possibili e ragionevoli eccezioni) e l'impiego di processi di cifratura (non si ha la garanzia di conservare opportunamente le chiavi per la decodifica essendo esse, per definizione, da mantenere segrete).
- Mantenimento della *fedeltà di presentazione del contenuto*: ciò non significa semplicemente conservare nel tempo la struttura del documento, ma anche i cosiddetti "font" e tutte le altre caratteristiche utili alla completa comprensione del documento.
- Da questa esigenza deriva anche il non dover contenere riferimenti ad altre informazioni o documenti (ad esempio link ipertestuali) non sotto il controllo della stessa "unità di informazione" o documento.
- Il formato di conservazione deve corrispondere fedelmente ad uno standard (formale o *de facto*) interamente descritto e le cui specifiche siano state rese di dominio pubblico senza alcuna forma di limitazione, presente o futura.
- Ciò garantisce, a fronte dell'imprevedibilità dell'evoluzione delle tecnologie digitali (sempre nell'assunto di poter rileggere fedelmente quanto registrato al momento dell'inizio del processo di conservazione), di essere in qualunque momento in grado di interpretare correttamente il formato interessato e di ricostruire fedelmente il suo contenuto informativo.
- Il processo di *conversione* dal formato di un documento operativo (in questo caso potrebbe essere anche un documento cartaceo) ad uno di conservazione non deve

obbligare a operazioni complesse ed impegnative dal punto di vista tecnico.

Ovviamente, la diffusione del formato (compresi i programmi di visualizzazione ad esso correlati) e la sua portabilità sono una garanzia non solo di qualità (maggiore diffusione equivale ad un maggior numero di test positivi eseguiti), ma anche di quella che alcuni definiscono *portabilità temporale*, ossia delle garanzie, insite nella esigenza che qualunque sistema ha di proteggere gli investimenti fatti, di analoga diffusione nel futuro di tale formato.

Alla data sono pochi i formati che soddisfano i requisiti appena illustrati. Essi sono:

- PDF/A-1, descritto nello standard ISO 19005-1, il quale ha anche il vantaggio di poter disporre di uno "strato" XML (in grado di trasportare dati strutturati e metadati) e la possibilità di "inglobare" in modo nativo un'eventuale firma digitale senza ricorrere a tecniche di imbustamento.
- Formato immagine (Raster) TIFF in versione non compressa o compressa senza perdita di informazione. Di questo formato vi sono parecchie versioni, alcune delle quali proprietarie (che ai fini della conservazione nel lungo periodo sarebbe bene evitare). In genere le specifiche sono pubbliche e non soggette ad alcuna forma di limitazione. Questi sono, inoltre, formati molto utilizzati per la conversione in digitale di documenti cartacei. *Il loro impiego, se attuano una forma di compressione, va valutato attentamente in funzione del tipo di documento da conservare* (ad esempio per i documenti di tipo commerciale può essere individuato un livello di compressione più che accettabile). Esistono, infine, alcuni formati ISO basati sulla specifica TIFF 6.0 di Adobe (che è quella "ufficiale" del TIFF).

Si tratta del formato ISO 12639, altrimenti noto come TIFF/IT, rivolto particolarmente al mondo del publishing e della stampa e dell'ISO 12234, altrimenti detto TIFF/EP, più orientato alla fotografia digitale.

- Il formato JPG è aderente alla norma ISO 10918-1, ma comporta una perdita di qualità dell'immagine originale. Anche in questo caso, come nel caso dei TIFF, avendo una grande diffusione, può essere considerato, ma *il suo impiego, correlato ad un opportuno livello di compressione, va valutato attentamente in funzione del tipo di documento da conservare.*
- Lo stesso gruppo che ha ideato il JPG ha prodotto il JPEG 2000 (ISO/IEC 15444-1) che può utilizzare la compressione senza perdita di informazione. Il formato JPEG 2000 consente, inoltre, di associare metadati ad un'immagine. Nonostante queste caratteristiche la sua diffusione è tutt'oggi relativamente limitata poiché è supportato solo da programmi di nuova generazione. In realtà è un formato complesso da sviluppare e pertanto il suo supporto non si è diffuso rapidamente come quello di altri formati malgrado le apprezzabili caratteristiche delle sue immagini a colori.

5. Alcune ulteriori considerazioni

Il fatto che non siano disponibili molti formati adatti alla conservazione nel lungo periodo del digitale *non deve apparire come un limite*, bensì va visto come positivo il fatto che ve ne siano: ciò, infatti, consente di avviare con una certa tranquillità processi di conservazione della memoria digitale, almeno per quanto riguarda le informazioni gestite in modo analogo ai documenti tradizionali. Altri formati sono in cantiere e probabilmente se-

guiranno a breve: l'importate è aver analizzato e compreso le problematiche sottese ed iniziare ad operare.

Un cenno meritano anche i formati ODF (*Open Document Format*), corrispondente allo standard ISO/IEC 26300, e Open XML di Microsoft (per il quale è in corso un processo di standardizzazione); questi formati offrono tutti i vantaggi derivanti dall'adozione di XML, ma non hanno le caratteristiche necessarie alla conservazione nel lungo periodo.

Sono infatti formati adottati da suite di tipo Office, quindi utilizzabili per attività di creazione e modifica dei documenti.

È comunque ancora presto per valutare le altre caratteristiche (vedi sopra) che ne possano consigliare l'impiego diffuso nel contesto operativo.

Se poi otterranno la diffusione sperata da chi li sostiene, non è da escludere che tali formati vengano dotati della possibilità di far loro acquisire, al momento in cui è necessario e su richiesta dell'utente, le caratteristiche proprie di un formato adatto alla conservazione nel lungo periodo.

Capitolo III

Il paradosso dell'informazione e del supporto su cui è registrata

Da quando gli esseri umani hanno iniziato a sentire l'esigenza di fissare su di un supporto fisico le loro esperienze, o i loro pensieri, si può parlare di comunicazione non più solo orale, ma in qualche modo *scritta*.

Di seguito, in modo puramente strumentale ai fini dell'esposizione, si ripercorre brevemente la storia della scrittura come mezzo di diffusione dell'informazione e il peso che a tali fini ha, e ha avuto, il supporto di registrazione dell'informazione stessa.

Oltre 40.000 anni fa sono apparse le iscrizioni rupestri, si è poi passati alla scrittura (3.000 – 4.000 anni a.C.), inizialmente incisa sulla pietra.

Queste iscrizioni sono tuttora leggibili a distanza di millenni, ma la possibilità della loro creazione e consultazione era limitata al luogo in cui si disponeva del supporto e la trasportabilità era praticamente nulla.

L'impiego delle tavolette di argilla, in seguito, ha consentito di liberarsi dal vincolo del luogo fisico in cui il supporto è disponibile per registrare e per rileggere le informazioni: le tavolette potevano essere scritte praticamente ovunque e trasportate in qualunque altro luogo per poter essere consultate. La relativa facilità con cui le tavolette potevano essere prodotte e scritte rappresentava, inoltre, un indiscusso potenziale di diffusione dell'informazione e pertanto anche una prima concreta possibilità di unire, integrare e completare informazioni, generando

nuova informazione che poteva poi essere registrata su nuovi supporti.

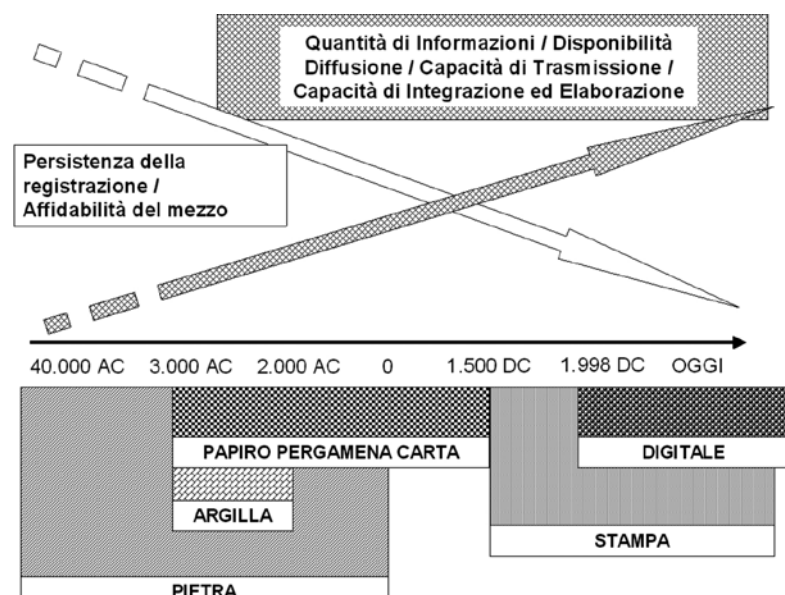


Figura 3 – Evoluzione delle caratteristiche dell'informazione in funzione del supporto sulla quale è registrata e trasmessa

L'intrinseca fragilità dell'argilla richiedeva, però, opportune cure nella loro conservazione. Ciò ha consentito che arrivassero sino ai giorni nostri tavolette di oltre 4.000 anni fa (una per tutte la scoperta, nel 1975, degli archivi reali di Ebla che hanno riportato alla luce oltre 17.000 tavolette d'argilla, con iscrizioni cuneiformi, risalenti al periodo tra il 2500 e il 2200 a.C.). Papi, pergamene e la stessa carta¹⁵ (fino all'invenzione della stampa

¹⁵ In occidente questa fu introdotta dagli Arabi, che ne avevano imparato l'arte dai Cinesi - che pare la conoscessero fin dall'inizio del II secolo - ed iniziarono a produrla poco prima dell'800 d.C.

a caratteri mobili¹⁶) hanno rappresentato da oltre due millenni prima di Cristo i principali strumenti per registrare, conservare e diffondere l'informazione.

Questi tipi di supporto (molto più facili da produrre e molto più maneggevoli e capienti dei precedenti) hanno aumentato la capacità di produzione di informazione, la sua diffusione e la possibilità di integrarla.

Il diffondersi dell'impiego di tali supporti ed il conseguente aumento dell'informazione prodotta hanno fatto emergere in modo sempre più marcato l'esigenza di conservare questi supporti, per non perdere e poter tramandare l'informazione in essi contenuta.

È proprio per soddisfare la succitata esigenza che si diffondono e si istituzionalizzano il concetto e la pratica di *copia* dei documenti, vuoi a tutto vantaggio della loro diffusione vuoi ai fini della loro conservazione.

A metà del XV secolo, grazie alla stampa a caratteri mobili, inizia il periodo della *carta stampata*: la facilità con cui si può passare da un manoscritto ad un numero praticamente illimitato di copie non solo aumenta la possibilità di produrre informazione, ma anche la sua diffusione e la possibilità di integrazione risultano estremamente facilitate.

Il problema della loro conservazione viene preso in carico da archivi e biblioteche che, oltre a catalogare opportunamente i documenti si preoccupano di conservarli in ambienti adatti. In queste sedi, ai fini conservativi, prendono forma e si affermano le operazioni di restauro, non solo dei documenti cartacei, ma anche di quelli delle epoche precedenti.

¹⁶ L'invenzione della stampa a caratteri mobili da parte di Johannes Gutenberg viene fatta risalire al 1439, ma è con la stampa della Bibbia, nel 1455, che si può considerare aver inizio l'era della stampa.

Si realizzano anche i primi esperimenti tecnologici ai fini conservativi e di diffusione della fruizione dei documenti: ad esempio è nel 1940 che iniziano ad affermarsi e diffondersi le tecnologie di microfilmatura (in molti casi tuttora in uso); negli ultimi vent'anni, principalmente ai fini di una maggiore facilità e diffusione della fruizione dell'informazione conservata, prendono poi piede tecnologie digitali quali quelle della gestione elettronica dei documenti.

Oggi la generazione dell'informazione avviene principalmente in formato digitale.

Essa deve essere registrata su supporti digitali e può essere rapidamente trasferita come tale¹⁷ - vuoi tramite tali supporti vuoi tramite reti - e raggiungere qualunque utente (essere umano o strumento dotato di potenza elaborativa) in qualunque parte del mondo; ciò consente anche possibilità di duplicazione e d'integrazione tra le informazioni (e pertanto la possibilità di produrre nuove informazioni) inimmaginabili solo qualche decennio addietro.

In conclusione, con il passare del tempo, con un'accelerazione sempre maggiore e grazie proprio agli strumenti usati per registrare le informazioni, si è avuta una sempre maggior facilità di produzione delle informazioni, una sempre maggior capacità di trasferirle, di renderle disponibili, di duplicarle e di integrarle. Di contro, la possibilità di perderle, o di veder alterata (per qualunque motivo) l'informazione originale, diventa sempre più probabile.

Oggigiorno le nostre informazioni sono, dunque, registrate e conservate su supporti digitali e non si può pensare, per la loro conservazione, di impiegarne altri.

¹⁷ Ovviamente non è escluso il trasferimento su supporto cartaceo.

Vale, pertanto, la pena di analizzare anche brevemente quali sono i supporti impiegati per registrare il digitale e di individuare quelli che, associati ad opportune tecnologie e metodologie, risultano oggi¹⁸ i più adatti a soddisfare le attuali esigenze di conservazione delle informazioni nel lungo periodo, di protezione contro la loro perdita o alterazione nonché di disponibilità per chi ne ha bisogno nel luogo e nel momento richiesto.

Va da sé che la componente supporto di registrazione è solo uno degli elementi che compongono un sistema di conservazione digitale; il compito di minima di questa componente è quello di garantire l'accesso alle registrazioni fisiche, ossia di consentire di rileggere correttamente i bit (così come erano stati originariamente registrati) da supporti che molto probabilmente saranno differenti da quelli originari.

Non ci si può a questo punto esimere dall'impiegare il termine *storage* (ormai entrato nell'uso comune) per indicare genericamente i singoli supporti di registrazione digitale o sistemi, più o meno strutturati, di questi ultimi.

Questa parola, tormentata e misteriosa, ha da sempre indicato, nei sistemi di calcolo, il *contenitore* dove sono memorizzati i dati che prima o poi saranno riutilizzati.

La semplicità del concetto è resa complessa da termini tecnici che a volte dividono anche gli addetti ai lavori, ma di fatto si tratta di contenitori di dati registrati su superfici, prevalentemente magnetiche, con caratteristiche tecniche differenzianti quali, ad esempio, la capacità di contenere dati e informazioni, la tipologia fisica del supporto di memorizzazione o le caratteristiche di accessibilità.

¹⁸ È implicito che in un futuro anche prossimo le tipologie di supporto, le tecnologie e le metodologie potranno evolversi anche in modo non prevedibile.

Capitolo IV

I supporti per registrare e conservare le informazioni digitali

1. Tecnologie ottiche: storage professionale e supporti di registrazione consumer, un confronto che dura da 25 anni

Dall'avvento del CD-R (*Compact Disc-Recordable*) e delle prime tecnologie storage ottiche professionali (seconda metà degli anni '80) si sono spesso operati, specialmente nel nostro Paese, confronti tra i due tipi di tecnologie.

È d'obbligo, innanzi tutto, una distinzione di base: le soluzioni professionali (gli ormai obsoleti dischi WORM¹⁹ da 12" e 14" e quelli da 5 1/4" ablativi²⁰ e magneto-ottici nonché gli attuali UDO, *Ultra Density Optical*) che si sono succedute nel tempo, sono state progettate e realizzate specificamente per l'archiviazione a lungo termine.

Al contrario, i CD, DVD, Blu-ray e HD-DVD registrabili, nelle loro molteplici versioni (e standard), sono stati pensati per applicazioni consumer di intrattenimento (audio, video, multimedia, videogiochi), ma grazie alle relativamente elevate capacità ed i bassi costi dei supporti, oltre a rappresentare anche strumenti utili per distribuire grandi quantità di informazioni a

¹⁹ Write Once, Read Many.

²⁰ La registrazione avveniva modificando il substrato di registrazione in modo permanente, operando da parte del raggio laser una scalfittura, in corrispondenza del segno 1. La differenza di potere riflettente tra 0 e 1 consentiva ad un raggio laser di minor potenza di rileggere correttamente quanto registrato.

costi convenienti, sono spesso diventati periferiche di storage *tout court* che, grazie ai dispositivi robotici (detti jukebox o librerie) hanno trovato posto nelle soluzioni di storage per la conservazione.

Una prima differenza importante è che i drive (dispositivi di lettura e scrittura) professionali sono espressamente progettati e realizzati per un ambiente automatizzato, mentre i drive consumer sono prodotti *stand-alone*: collegabili sì internamente o esternamente ad un PC, ma che devono essere profondamente modificati per una eventuale integrazione in un jukebox.

I dischi professionali sono, poi, architetturelmente protetti da una cartuccia, mentre quelli consumer sono dischi “nudi”.

Questa differenza è fondamentale e si ripercuote sia sul costo e sulla qualità dei supporti, sia sulla loro gestione, durata, affidabilità e longevità.

I dischi consumer sono quindi più a rischio nei confronti della loro gestione fisica e della contaminazione ambientale, diminuendo ancora la loro già modesta capacità di durata nel tempo che nella migliore delle ipotesi è data sui 10 anni, contro gli almeno 50 anni dei supporti ottici professionali.

Va anche ricordato che, proprio per la loro diffusione sul mercato, i supporti vergini per la masterizzazione (registrazione) dei prodotti consumer disponibili possono variare drasticamente in termini di materiali, qualità e prezzo.

Dal punto di vista architetturel è importante anche evidenziare che i prodotti consumer non prevedono la rilettura di una registrazione appena effettuata per verificarne la correttezza, né dispongono di sistemi di controllo per superare e tener traccia di eventuali errori di lettura.

I prodotti e le architetture professionali tendono, poi, ad avere un'evoluzione ed una crescita che garantiscano la protezione degli investimenti del cliente/utente.

I supporti consumer, invece, seguono dinamiche di mercato e privilegiano scelte di standard, compatibilità ed evoluzione completamente diverse e con prospettive temporali molto più limitate; a titolo di esempio si veda l'ormai superata contesa per l'affermazione sul mercato della videoregistrazione tra VHS e Betamax. Non è uscito vincitore né il prodotto migliore né la cordata più forte, ma piuttosto ciò che il mercato ha percepito avere il miglior rapporto prezzo/prestazioni per le esigenze predominanti.

Un analogo scenario si presenta oggi: chi la spunterà sul mercato tra Blu-ray e HD-DVD? È davvero difficile dirlo, di certo la tecnologia che non otterrà il predominio è destinata a svanire nel corso di qualche anno, o poco più, a danno di chi la avrà scelta.

Esiste, infine, un piccolo accorgimento per individuare, senza incertezze, se un disco ottico è un prodotto professionale o consumer: i produttori di dischi ottici professionali chiamano i propri prodotti “disk”, analogamente agli hard disk (MO disk, UDO disk, ecc.). I prodotti consumer invece della desinenza “k” usano la “c” (CD-R disc, DVD disc, Blu-ray disc). È una piccola differenza ortografica che dichiara subito una profonda diversità architetturel!

2. Le tecnologie storage disponibili

Concentrandoci sulle tecnologie professionali si hanno attualmente tre grandi categorie: i dischi magnetici, i nastri magnetici e i dischi ottici.

Ognuno di tali dispositivi/sistemi ha i suoi pregi ed i suoi difetti e va a soddisfare ben precise esigenze di registrazione e conservazione.

2.1 Dispositivi a disco

La tecnologia di base, utilizzata attualmente nei dispositivi a disco, si fonda sulla registrazione indotta da un campo magnetico che, per determinati materiali, avviene in modo permanente.

Questa permanenza rende possibile l'esistenza dello storage non-volatile.

Il dispositivo più diffuso nel panorama dei dispositivi di memorizzazione è il disk drive magnetico, in gergo *hard disk* o *hard disk drive* (HDD).

1 Kilobyte = 1.000 byte
1 Megabyte = 1.000 Kilobyte
1 Gigabyte = 1.000 Megabyte
1 Terabyte = 1.000 Gigabyte
1 Petabyte = 1.000 Terabyte
1 Exabyte = 1.000 Petabyte
1 Zettabyte = 1.000 Exabyte
1 Yottabyte = 1.000 Zettabyte

Figura 4 - Le unità di misura della capacità dello storage

(Ormai è invalso l'uso di non utilizzare il sistema binario per indicare la quantità di informazione da registrare ovvero la capacità di un supporto di memorizzazione, bensì di impiegare il sistema di misura decimale)

Solitamente è composto da uno o più dischi in alluminio, o vetro, rivestiti di uno strato magnetico in rapida rotazione e da due testine per ogni disco (una per lato), le quali, durante il funzionamento “volano” alla distanza di pochi centesimi di nanometro dalla superficie del disco leggendo e scrivendo i dati. La rota-

zione dei dischi varia tra i 4.500 e i 15.000 giri al minuto e ciascuna faccia, o lato del disco, ha una propria testina di lettura e registrazione dei dati.

Di dischi magnetici ne esistono per tutte le esigenze: a livello di prestazioni, costi, capacità, scalabilità, consumo, modalità di connessione a reti e sistemi di elaborazione.

Le caratteristiche principali di un hard disk moderno sono: la capacità, il tempo di accesso e l'affidabilità.

La *capacità* è in genere espressa in gigabyte (GB). Attualmente gli hard disk disponibili hanno capacità comprese tra 40 gigabyte e 1 terabyte; la capacità può essere aumentata incrementando la densità con cui le informazioni sono memorizzate sui dischi, oppure usando dischi più grandi o impiegandone un numero maggiore. La rincorsa a unità sempre più capienti non conosce soste.

Il *tempo di accesso* è la variabile più importante nel determinare le prestazioni di un hard disk. Si tratta del tempo medio necessario perché un dato, residente in un punto casuale dell'hard disk, possa essere reperito. Il tempo impiegato dipende dalla velocità della testina a spostarsi sulla traccia dove risiede il dato e dalla velocità di rotazione del disco: maggiore è quest'ultima e più breve è il tempo impiegato dal dato a ripassare sotto la testina nel caso questa non fosse arrivata in tempo sul dato, durante la rotazione precedente. Il tempo di accesso tipico per un hard disk ad alte prestazioni (15.000 giri) è di 3 o 4 millisecondi.

L'*affidabilità* viene normalmente espressa dal *Mean Time Between Failure* (letteralmente: tempo medio fra i guasti), spesso abbreviato in MTBF. È questo un parametro di qualità applicabile a dispositivi meccanici, elettrici ed elettronici e ad applicazioni software. Il MTBF indica il valore atteso del tempo tra un guasto ed il successivo.

L'MTBF tipico per un hard disk ad alte prestazioni è di 800.000 ore, equivalenti a novantuno anni, un valore sufficiente a tranquillizzare qualsiasi utente, ma il numero di dischi e la criticità dei dati rende necessario, a garanzia della non perdita di dati, l'utilizzo di tecniche addizionali quale il RAID.

2.2 I Sistemi RAID

Un *Redundant Array of Independent Disks* (insieme ridondante di dischi indipendenti, RAID), originariamente *Redundant Array of Inexpensive Disks*, è un sistema che usa un insieme di dischi rigidi per condividere, o replicare, le informazioni.

I benefici di un sistema RAID sono di aumentare l'integrità dei dati, la tolleranza ai guasti e/o le prestazioni, rispetto all'uso di un disco singolo.

Le sue prime specifiche furono sviluppate presso l'Università di California a Berkeley: esse definiscono le modalità di registrazione di dati in più punti (registrazioni ridondanti) di un sistema con più dischi fisici (supporti, testine, controller).

Al suo livello più semplice, un sistema RAID permette di combinare un insieme di dischi in una sola unità logica; in questo modo il sistema operativo, invece di vedere differenti dischi, ne vede solamente uno.

Un blocco in un sistema a dischi è spesso composto da più settori di disco. Una serie di blocchi consecutivi è detto *stripe* (= striscia).

Un sistema RAID è visto dal sistema operativo come un unico disco logico, mentre la tecnica di registrazione (striping) prevede la suddivisione di ogni disco fisico in "partizioni" (stripe) che possono avere le dimensioni minime di un settore (512 byte) ed

arrivare fino a più megabyte.

Il risultato di registrare dati (partizioni) su più dischi può avere come risultato:

- aumentare le prestazioni, se le operazioni di I/O possono essere opportunamente sovrapposte e bilanciate;
- aumentare l'affidabilità del sistema, in quanto si aumenta il *Mean Time Between Failure* (MTBF).

Per quanto concerne le prestazioni, ad esempio, un sistema RAID, essendo composto da più dischi fisici, consentirà di disporre di una capacità massima equivalente alla somma dei singoli dischi.

Inoltre, la possibilità di accedere in parallelo ai singoli dischi aumenta di un fattore eguale al parallelismo consentito, la velocità di trasferimento dei dati.

Dal punto di vista dell'affidabilità si possono considerare due situazioni:

- mirroring – ogni registrazione è duplicata simultaneamente su due dischi fisici differenti. È una soluzione costosa in quanto richiede un numero doppio di dischi fisici, ma è molto efficace poiché la rottura simultanea dei due dischi è molto improbabile; una registrazione fa, dunque, da back-up all'altra;
- controllo di parità – questo metodo oltre a suddividere l'informazione registrata su più dischi, registra anche il risultato di un algoritmo (*checksum*) che consente di ricalcolare una porzione di informazione se questa viene persa.

Vi sono innumerevoli modi, anche proprietari, di realizzare un Sistema RAID; questi rappresentano una serie di compromessi

tra prestazioni e affidabilità; l'aumento delle prestazioni e l'affidabilità (ridondanza) sono in opposizione: se si desidera ridondanza di dati si deve perdere in prestazioni e viceversa.

I due casi estremi sono:

- RAID-0. Consente il massimo livello di prestazioni in quanto le registrazioni non avvengono registrando in sequenza tutti i bit, ma sono eseguite in parallelo nel numero massimo di partizioni (dischi differenti) disponibili. In tale situazione si ha il minimo livello di affidabilità possibile.
- RAID-1. Si tratta di *Mirroring*. I dati sono duplicati su almeno due dischi fisici. Le prestazioni in scrittura non variano da quelle di un solo disco, mentre quelle in lettura possono trarre vantaggio dalla lettura simultanea del numero di copie eseguite. L'affidabilità è massima. L'esigenza di garantire sempre maggior affidabilità e prestazioni ai sistemi a disco ha portato all'introduzione nei sistemi RAID di dischi spare, da un minimo di uno fino ad un massimo a discrezione dell'utente, da utilizzarsi automaticamente in caso di cedimento per sostituirsi ad un altro disco che per qualunque motivo dovesse guastarsi.

2.3 Tipologie di dispositivi a disco

I dispositivi a disco disponibili sul mercato, come è facile immaginare, hanno caratteristiche fisiche e di prestazioni differenti tra loro in funzione dell'impiego cui sono destinati, del produttore, ecc.

Possiamo raggrupparli genericamente in quattro categorie: desktop disk, enterprise disk, mobile disk e consumer disk.

Tralasciando le ultime due categorie, è utile fare qualche consi-

derazione sulle differenze principali tra i dispositivi a disco della categoria desktop e quelli della categoria enterprise.

I primi sono parte del processo di concorrenza, molto forte peraltro, di un mercato di tipo commodity e subiscono una continua diminuzione del prezzo comparato alla capacità offerta.

I dischi della categoria enterprise, invece, hanno un ciclo di vendita più lungo e, in virtù di una richiesta applicativa tipica degli ambienti aziendali, privilegiano la *densità di accesso*²¹ invece del prezzo in relazione alla capacità in spazio.

Questo fa sì che la ricerca sia rivolta verso una minor dimensione del supporto ed una maggiore velocità di rotazione del disco stesso.

In questo modo viene soddisfatta anche un'esigenza, molto sentita in questi ultimi tempi, come quella del risparmio energetico: minori sono le dimensioni fisiche del disco e migliore è il tempo di accesso al dato senza necessariamente spendere più energia per raggiungere il dato stesso.

L'esigenza di immagazzinare una sempre maggior quantità di informazioni digitali, magari di non rilevante criticità per l'organizzazione, ma che devono comunque essere gestite introduce, infine, sul mercato tecniche che consentono di unire più dischi magnetici al fine di poter disporre di "contenitori" di dimensioni sempre maggiori.

²¹ È un modo per misurare le prestazioni di un disco nel suo complesso. L'incremento esponenziale della capacità dei dischi magnetici ha sì rappresentato per l'utente una forte riduzione nei costi per unità di memorizzazione, ma parallelamente non sono cresciuti in proporzione gli altri elementi che contribuiscono alle prestazioni con il risultato di doversi attendere prestazioni inferiori o rimediare con compensazioni quali maggior *caching* (attività di lettura da disco fisico e messa a disposizione di dati, che presumibilmente saranno richiesti, su una memoria ad accesso molto veloce), meccanismi di controllo più sofisticati, ecc.

Di seguito due delle principali tecnologie in proposito:

- *JBOD* (Just a Bunch of Disks). È una concatenazione tra dischi magnetici che consente di unire un insieme di dischi fisici in un grande disco virtuale al fine di far sembrare l'insieme come un singolo disco. JBOD usa due o più dischi fisici per creare un singolo disco logico. Anche il JBOD, come il RAID 0, nel caso di rottura di un disco rende inutilizzabile l'intero array di dischi. In questo caso però il recupero dei dati sarà un po' più probabile rispetto al RAID 0 in quanto i dati sono scritti linearmente come se fosse un solo hard disk con un numero di blocchi danneggiati pari alla dimensione dell'hard disk che si è rotto.
- *MAID* (Massive Array of Idle Disks). È una tecnologia che impiega un gran numero di dischi fisici; di questi sono attivati solo quelli in uso. Ciò riduce il consumo di energia (alimentazione e raffreddamento) e allunga la vita dell'hardware. Un MAID può essere composto da centinaia o migliaia di singoli dischi. Un tale sistema consente di disporre di memorie di massa ad un costo per terabyte abbastanza vicino a quello dei nastri. Un sistema MAID è, normalmente, costituito da dischi SATA a basso costo, consentendo di contenere molti più dati di un sistema RAID di eguale costo, con esigenze di alimentazione e raffreddamento di molto inferiori. È, inoltre, più voluminoso ed ha prestazioni inferiori (*throughput* inferiori, tempi di latenza più lunghi, minori ridondanze, ...).

2.4 Le interfacce tra disco e host

Host è un termine tecnico che indica il “sistema” composto da processore, sistema di calcolo e sistema operativo; l'interfaccia

di comunicazione ha il compito di permettere l'utilizzo dei dati che risiedono sullo storage da parte del sistema e di trasferire i dati elaborati dal sistema allo storage.

Uno storage è accessibile da più sistemi di calcolo con diversi sistemi operativi. Ci sono molte tipologie di interfacce.

Un'interfaccia può essere pensata schematicamente come l'insieme (combinazione) formato da:

- il *materiale* della connessione fisica, che può essere un filo di rame o una fibra ottica;
- il *tipo* di connessione fisica utilizzato:
 - parallel SCSI: un fascio di cavi con un connettore largo e piatto,
 - IDE o PATA²²: un nastro di cavi piatto e solitamente connesso internamente ad un PC;
 - SATA²³ e SAS (*Serial Attached SCSI*): utilizzano la stessa tipologia di cavo a 4 fili;
 - Fibre Channel: connessione a fibre ottiche (esiste anche la versione con rame, ma viene utilizzata solo all'interno dell'array disco);
 - FATA (*Fibre Channel ATA*): connessione fibre channel a fibre ottiche;
 - USB (*Universal Serial Bus*) e FireWire/IEEE 1394: tipologia di connessione plug and play per connessioni esterne che possono essere di dischi ATA o SATA;

²² ATA (Advanced Technology Attachment) è un'interfaccia standard per la connessione di dispositivi di memorizzazione (hard disk, CD/DVD, ...) in un personal computer. Sono molti i termini utilizzati per designare tale standard, comprese abbreviazioni e acronimi quali IDE, EIDE, ATAPI, UDMA e Parallel ATA (PATA).

²³ SATA, Serial attached ATA, è una tecnologia impiegata nella realizzazione di dischi magnetici di grande capacità a costi contenuti. Il Serial ATA è l'evoluzione dell'ATA.

- il *protocollo* o il linguaggio utilizzato e “parlato” nelle connessioni per la lettura o la scrittura del dato stesso
 - SCSI: viene usato per descrivere sia la connessione sia il protocollo del linguaggio (anche la connessione FC usa il medesimo protocollo);
 - ATA (deriva da PC IBM AT attachment) e solitamente indica un protocollo semplificato nel “command set” dei dischi.

Per ragioni storiche e di mercato, il segmento enterprise utilizza il protocollo SCSI mentre il mercato dei dischi a basso costo ed alta capacità impiega principalmente la tecnologia ATA.

Ci sono altre ragioni per cui il mercato enterprise preferisce SCSI e sono da ricercare nella sua affidabilità dovuta all’uso di componenti più costosi.

2.5 Le tecnologia delle memorie a nastro magnetico digitale

La memoria digitale a nastro magnetico è un dispositivo storage ad accesso sequenziale utilizzato dai computer fin dal lontano 1952.

Inizialmente il supporto a nastro era avvolto su una bobina a vista ed era montato manualmente nelle unità di lettura/scrittura. Negli anni ’80, per proteggere meglio il supporto dagli agenti esterni, il nastro viene racchiuso in cartucce sigillate.

Alla fine degli anni ’80 compaiono le librerie nastro automatizzate che ospitano all’interno un gran numero di cartucce e consentono il loro caricamento nei dispositivi di lettura e scrittura in modo completamente automatico.

Oggi è disponibile sul mercato un’ampia gamma di librerie au-

tomatiche con capacità interna a partire da poche decine fino a diverse migliaia di cartucce.

La capacità di memorizzazione del nastro si è evoluta tantissimo: una bobina nel 1956 conteneva 1,4 MB, una delle prime cartucce degli anni ’80 200 MB e nel 2007 sono state rese disponibili cartucce da 800 GB.

Di pari passo è cresciuta anche la velocità con cui è possibile leggere o scrivere dati su nastro: nel 1956 era di soli 7,5 KB/sec mentre oggi arriva a 120 MB/sec.

Per la loro natura di dispositivo ad accesso sequenziale i nastri magnetici sono oggi ampiamente utilizzati nei data center per realizzare operazioni di back-up di massa a basso costo grazie all’elevata capacità di trasferimento dati che oggi con drive LTO4 arriva fino a 846 GB/ora.

I nastri magnetici sono anche ampiamente utilizzati per l’archiviazione di dati storici, per dati con accesso poco frequente, per la conservazione a norma e di lungo periodo.

2.6 Lo standard LTO

Nel 1997, al fine di superare la proliferazione di formati e di tecnologie incompatibili tra loro, garantire l’interoperabilità multi-fornitore delle cartucce, assicurare la compatibilità all’indietro (lettura/scrittura di cartucce delle generazioni precedenti con la generazione corrente) e pianificare l’evoluzione in termini di capacità, prestazioni ed affidabilità, HP, IBM e Seagate (ora Quantum) propongono un nuovo *open standard* di cartuccia chiamato LTO (Linear Tape Open), oggi gestito dal Consorzio LTO²⁴.

²⁴ Sito ufficiale del consorzio LTO: www.ultrium.com.

Attualmente lo standard LTO è supportato da circa 30 produttori e dal 2007 sono sul mercato le unità e le cartucce LTO di quarta generazione con capacità nominale, prima della compressione, di 800 GB.

Una caratteristica comune a tutte le unità nastro di tipo professionale è la funzionalità integrata di compressione dati: nel caso di LTO 4 con un rapporto di compressione medio di 2:1 la capacità arriva a 1,6 TB.

A dieci anni dalla sua nascita, lo standard LTO ha superato la quota di 80 milioni di cartucce e di 1,5 milioni di drive spediti a conferma della validità della iniziativa di standardizzazione proposta.

2.7 Affidabilità e durata nel tempo

Dal punto di vista dell'affidabilità vale la pena di accennare ai seguenti fattori.

Le informazioni sono scritte a blocchi corredati da codici per la correzione degli errori e i dati, dopo essere scritti, sono *riletti per verificarne l'accuratezza*.

Le cartucce sono dotate di un chip di memoria che contiene una registrazione di tutte le segnalazioni di errore e consente di tracciare la qualità del supporto e provvedere ad eventuali *azioni preventive per evitare la perdita di informazioni*.

Grazie a questi accorgimenti, una cartuccia nastro di buona qualità, se conservata in condizioni ambientali controllate, può mantenere inalterato il contenuto fino a 25-30 anni.

L'evoluzione tecnologica porta alla disponibilità di supporti sempre più capaci (raddoppio della capacità ogni 18-24 mesi),

più veloci e più affidabili, e consente di ridurre nel tempo il costo di memorizzazione per unità di memoria.

Grazie alla diffusione delle librerie a nastro il processo di riversamento è oggi totalmente automatizzato: dal punto di vista operativo il gestore del data center alimenta la libreria con nuove cartucce e ritira alla fine quelle scartate dal processo. Per decidere quali non sono più idonee a far parte dell'archivio si analizzano le occorrenze di errore registrate nel chip integrato e, in base a limiti stabiliti dal gestore, quando una cartuccia raggiunge una soglia di errori prefissata viene attivata una operazione di riversamento e la cartuccia è successivamente scartata.

La gestione del catalogo delle cartucce, del riversamento periodico, della migrazione a nuove tecnologie e delle copie multiple per la protezione dati è di solito svolta dal software di gestione dello storage sulla base di politiche di gestione specifiche in termini di durata della archiviazione e di esigenze di protezione per ogni categoria di dati; la gestione avviene in modo automatico sfruttando tutte le risorse disponibili.

Per poter indirizzare specifiche esigenze di conformità di archiviazione sono disponibili anche speciali cartucce di tipo WORM: particolari accorgimenti nella formattazione del supporto e nella realizzazione fisica della cartuccia rendono impossibile alterare i dati una volta scritti.

Per favorire la protezione di dati sensibili, ed essendo le cartucce facilmente trasportabili, è stata resa disponibile in modo integrato la possibilità di cifratura dei dati. Questa funzionalità è attivabile ogni volta che si registra una nuova cartuccia. In questo modo si possono proteggere le cartucce in transito (ad esempio: scambio dati tra enti o cartucce di disaster recovery) e prevenire il furto d'informazioni.

Negli ultimi anni sono state sviluppate anche soluzioni di librerie nastro *virtuali*.

In sostanza l'idea è quella di utilizzare un insieme di dischi magnetici ad accesso diretto per simulare la disponibilità di un gran numero di cartucce e unità di lettura e scrittura virtuali. Ciò consente di attivare diverse attività di scrittura e/o lettura in parallelo sui molteplici dispositivi virtuali.

Le librerie virtuali sono utilizzate prevalentemente quando si vuole parallelizzare al massimo le operazioni di lettura e scrittura nei salvataggi e quando si ha l'esigenza di ripristinare dati di piccole dimensioni senza richiedere il caricamento di una cartuccia e la ricerca sequenziale dell'oggetto.

In un progetto di archiviazione queste librerie possono essere utilizzate come memoria di livello superiore a quello dei nastri con un tempo di accesso più veloce, ma a costi superiori e quindi in ultima analisi per contenere i dati più frequentemente utilizzati dove è richiesto un livello di servizio più alto.

2.8 I dischi ottici UDO

La tecnologia professionale UDO (dischi ottici amovibili e relative librerie) è sul mercato dal 2003 ed oggi è disponibile nella sua seconda generazione da 60 GB per supporto. È stata la prima tecnologia a laser blu disponibile sul mercato e a tutt'oggi è quella con le caratteristiche più avanzate.

È una tecnologia professionale, specificamente progettata e realizzata per una strategia di conservazione a lungo termine di dati.

Le librerie sono disponibili con capacità da 24 a 638 dischi (da 720 GB a 38 TB). Per consentire un accesso veloce ai dati è stato

prodotto un *Archive Appliance* (soluzione NAS, dotata di cache di dischi magnetici da 2 TB) con capacità fino a 76 TB.

I dischi UDO hanno una ben precisa roadmap che prevede un raddoppio della capacità ogni 2-3 anni ed un significativo miglioramento delle prestazioni ad ogni cambio generazionale. Si prevede che nel 2009 la capacità per supporto sarà di 120 GB e nel 2012 di 240 GB, mantenendo totale compatibilità verso le generazioni di prodotto precedenti.

I supporti UDO sono protetti da una cartuccia che salvaguarda i dati da ogni possibile interferenza esterna: luce, temperatura, umidità, contaminazioni. Chi le produce le garantisce per 50 anni, anche se i test di invecchiamento artificiale ne evidenziano una durata e stabilità molto superiore, nell'ordine dei 180 anni.

L'UDO è inoltre disponibile nella versione *True WORM*.

Uno dei vantaggi pratici della tecnologia UDO è la longevità del supporto che consente di evitare migrazioni di dati non necessarie per periodi molto lunghi, anche 15-18 anni, rispetto ai cicli di obsolescenza tecnologica di altre tecnologie.

Un altro aspetto in cui l'UDO può vantare una supremazia sulle altre tecnologie è il consumo energetico, che è di circa 38 volte inferiore rispetto ad equivalenti sistemi a disco.

Per completezza d'informazione va evidenziato che recentemente (gennaio 2009) tutti gli asset (inclusi i brevetti, le tecnologie e gli stabilimenti produttivi) di Plasmon, unico produttore di queste tecnologie, sono stati acquisiti da Alliance Storage Technologies che ha dichiarato di voler continuare a supportare ed investire in queste tecnologie.

3. Architetture per lo storage in rete

Considerare la tecnologia dei singoli sistemi storage è, però, riduttivo in quanto le capacità di collegamento e la velocità di trasmissione delle reti hanno consentito di connettere i vari sistemi storage professionali in sistemi di rete ben più ampi, in grado di vederli e gestirli secondo le più diverse architetture, modalità di collegamento e protocolli (in pratica tutti quelli disponibili). I dispositivi di storage si sono, quindi, avvantaggiati delle architetture di rete per rendersi disponibili a connessioni multiple con altri dispositivi di storage e per far condividere dati tra più sistemi di calcolo.

In poche parole le modalità di collegamento tra dispositivi storage e sistemi si possono suddividere in:

- DAS, Direct Attach Storage, è un termine utilizzato per descrivere i dispositivi di storage che non sono inclusi in una rete e che sono direttamente connessi ai sistemi di calcolo.
- NAS, Network Attached Storage, si riferisce in modo esclusivo a dispositivi di storage che sono connessi ad una rete IP standard e che operano (consentono l'accesso) a livello di file e non di blocchi.
- SAN, Storage Area Network, è un'architettura che collega i dispositivi di storage ad una rete/area esclusiva ed ottimizzata per questi dispositivi (consentendo l'accesso a livello di blocchi) e utilizza un protocollo di comunicazione SCSI.

3.1 Infrastrutture di storage scalabili (Grid Storage)

Con la crescente disponibilità di collegamenti e la sempre maggior velocità delle reti si è ottenuto che i dispositivi di storage

possano colloquiare facilmente tra di loro e con i sistemi a loro collegati.

Parallelamente è cresciuta l'esigenza per le organizzazioni di disporre di sempre maggior spazio per memorizzare le informazioni che, come si è visto, crescono in modo veramente impressionante.

È facile comprendere che oltre un certo limite, la semplice crescita della capacità dei dispositivi non è sufficiente.

Il concetto stesso di scalabilità, legata ai dispositivi, ha come base l'esigenza di aumentare non solo la capacità fisica, ma parallelamente anche le prestazioni; entrambe devono essere aggiunte a quelle esistenti in maniera graduale ed a costi contenuti.

Si stanno quindi affermando le tecnologie *Grid Storage*.

Con questo termine si fa riferimento ad una modalità di memorizzazione dei dati che si basa su più nodi storage autoconsistenti (dotati, cioè, oltre che di memoria di archiviazione, di un microprocessore e di capacità di indicizzazione e di gestione autonomi), e interconnessi in modo tale che ognuno possa comunicare con qualunque altro senza dover passare per un "commutatore" centrale.

Questo tipo di tecnologia comporta almeno tre vantaggi aggiuntivi:

1. introduce un ulteriore livello di ridondanza e di possibilità di superare eventuali situazioni di errore (fault-tolerance) - se un nodo ha un malfunzionamento lo si può cortocircuitare con percorsi alternativi;
2. l'esistenza di più percorsi tra nodi garantisce stabilità di prestazioni anche in condizioni di carichi non costanti;
3. è facilmente e gradualmente scalabile: se si aggiunge un

nuovo nodo, esso può essere automaticamente riconosciuto dagli altri, con evidenti vantaggi anche in termini di fermo macchina.

Le funzionalità operative, personalizzabili, che ogni singolo nodo è chiamato a svolgere sono, inoltre, flessibili e determinate dal software che in esso viene caricato.

Queste tipologie di approccio, infine, si prospettano sufficientemente economiche in quanto prevedono:

- per i nodi elaborativi della struttura l'utilizzo di server a basso costo con dischi interni di tipo DAS;
- una rete economica e ad alte prestazioni per collegare tra loro i blocchi e i blocchi al mondo esterno;
- un software che presenti la struttura come un unico dispositivo di storage e sia in grado di fare da interfaccia alle richieste di ricerca delle informazioni e del loro aggiornamento.

4. La virtualizzazione

L'esigenza di disponibilità nel tempo del dato tende ormai, in molte situazioni, a raggiungere lassi di tempo sempre maggiori (come minimo un decennio); il ciclo di vita dell'infrastruttura tecnologica preposta alla sua conservazione è, in confronto, molto breve: le singole risorse hardware IT hanno un ciclo di vita che oscilla tra i 3 e i 10 anni; le soluzioni a disco hanno una durata che varia tra i 3 ed i 5 anni, i sistemi basati su nastro tra i 5 ed i 15 - 20 anni.

Questo aspetto merita la giusta attenzione se si pensa che l'accessibilità al dato è importante tanto quanto il dato stesso.

Il problema da risolvere è pertanto come *evitare continue migra-*

zioni e allungare per un tempo "illimitato" il ciclo di vita della risorsa storage.

La realtà, dal canto suo, si presenta con aziende che si sono dotate di apparati storage eterogenei, ovvero di sistemi acquisiti da fornitori diversi e di diversa tipologia.

Questo fenomeno non è solo frutto di scarsa pianificazione, ma anche la conseguenza di reali esigenze operative: le informazioni presenti all'interno di un'azienda sono di diversa natura e possono richiedere di essere gestite in modo differente, associando risorse storage differenti, non solo per ottenere i dovuti equilibri economici, ma soprattutto per garantire le giuste prestazioni.

Avere sistemi eterogenei vuol dire, oltre ad avere una maggiore complessità di gestione, non sfruttare appieno le funzionalità disponibili per proteggere e rendere accessibili le informazioni.

La soluzione al succitato problema consiste, dunque, nel trovare un modo per *poter considerare sistemi (storage) diversi come un'unica risorsa e quindi renderli interoperabili!*

Una prima possibile soluzione potrebbe essere quella di affidarsi ad un unico fornitore, ma in questo modo si rimane legati alle sue scelte tecnologiche e strategiche, nonché all'andamento dei suoi investimenti in ricerca e sviluppo.

In pratica, ci si può affidare a soluzioni di *virtualizzazione*, nelle varie realizzazioni.

Queste permettono di utilizzare storage differenti e di renderli visibili al server e/o agli applicativi come un'unica risorsa, indipendentemente dal modo in cui sono utilizzati fisicamente.

Secondo Wikipedia, "per virtualizzazione si intende la creazione di una versione virtuale di una risorsa normalmente disponibile

ed utilizzata fisicamente. Qualunque risorsa hardware o software può essere virtualizzata: sistemi operativi, server, memoria, spazio disco, sottosistemi. Un tipico esempio di virtualizzazione è la divisione di un disco fisso in partizioni logiche.

Meccanismi più avanzati di virtualizzazione consentono la ridefinizione dinamica sia delle caratteristiche della risorsa virtuale, sia della sua mappatura su risorse reali.

La virtualizzazione permette l'ottimizzazione delle risorse e la capacità di far fronte a esigenze specifiche secondo il più classico paradigma dell'*'on demand'*."

La virtualizzazione nello storage esiste, di fatto, da molto tempo, anche se non è mai stata chiamata con questo nome; disco virtuale, dischi logici, logical unit (LUN), volumi logici sono già una forma di virtualizzazione.

Il vero limite di queste tecnologie è il perimetro in cui agiscono, ossia all'interno del singolo sistema, e la migrazione dei dati tra un sistema e l'altro quando sopraggiunge l'obsolescenza.

Oggi la tecnologia di virtualizzazione ha raggiunto un elevato livello di maturità proprio perché *riesce a raggruppare diversi sistemi eterogenei in un unico pool*.

La principale differenza, infatti, tra un volume virtuale e una LUN è che il volume virtuale può essere creato, modificato e presentato in modi diversi, indipendentemente dal sottosistema storage in cui risiede.

Il volume virtuale può essere costituito da spazio storage allocato su differenti sistemi di storage, con diverse caratteristiche, e lo spazio dedicato può essere incrementato mediante qualsiasi storage disponibile. La configurazione di una LUN è, invece, legata al singolo sottosistema.

Possiamo, quindi, identificare nella virtualizzazione la capacità di mascherare la realtà fisica di un oggetto, o di un insieme di essi, e mostrarli come un'unica entità logica in grado di integrare i dispositivi fisici con procedure standard. L'idea di base è quella di trasformare i singoli sistemi in un unico pool che venga visto come un unico sistema, dove un livello software si occupa di tradurre le richieste applicative, o del sistema operativo, ai singoli dispositivi.

Ciò risolve anche il problema di rendersi indipendenti dai limiti temporali dettati dal ciclo di vita dei singoli sistemi di storage.

Più in generale i principali vantaggi di una soluzione di virtualizzazione sono:

- possibilità di avere una soluzione integrata di virtualizzazione trasversale a tutto il datacenter (server, storage, software e networking);
- risparmio economico;
- maggior semplicità di gestione in ambienti complessi;
- possibilità di allocare risorse IT non solo economiche, ma anche fisiche ed umane, su progetti diversi;
- miglioramento del livello di servizio e necessità di tempi ridotti per la manutenzione;
- *totale indipendenza dai fornitori* e dalla piattaforma stessa;
- possibilità di definire un piano di business continuity e disaster recovery comune a piattaforme ed ambienti eterogenei;
- possibilità di utilizzare ambienti storage eterogenei utilizzando le stesse funzionalità gestionali (copie locali, remote, ecc.);
- aumento del numero di sistemi gestibili dal singolo amministratore di sistema;

- limitate operazioni di migrazione;
- sfruttamento razionale dello spazio su disco e nastri;
- opportunità di utilizzare tecnologie di interconnessione avanzate, ma poco costose.

Va ricordato che sono disponibili soluzioni di virtualizzazione adottabili, a costi contenuti, da aziende di modeste dimensioni: anch'esse hanno le stesse esigenze di conservazione e disponibilità dei dati delle società di più grandi dimensioni.

Una soluzione di virtualizzazione in queste situazioni ha l'ulteriore vantaggio di permettere di continuare a sfruttare dispositivi di tecnologia matura già in uso.

4.1 I livelli di virtualizzazione

Si possono distinguere almeno tre diversi livelli architetturali in cui può essere realizzata la virtualizzazione; a livello host, a livello del sottosistema storage e a livello di network.

La virtualizzazione *a livello host* consiste nell'installazione su uno o più server di un software di gestione dello storage che permette a più sottosistemi di lavorare in parallelo con molteplici server. La virtualizzazione può essere però realizzata solo su uno storage già assegnato, senza la possibilità di ottenere una completa indipendenza tra i volumi e i server.

La virtualizzazione *a livello del sottosistema di storage* rappresenta una delle soluzioni più comuni di virtualizzazione in uso oggi.

Con questo metodo si creano volumi virtuali con lo spazio storage di uno o più sottosistemi, includendo la parte software nel controller. In tal modo viene garantita l'indipendenza dell'host.

Per realizzare una virtualizzazione *a livello di network*, infine, si inserisce un software nello switch di una SAN. Questo software utilizza propri metadati per creare e gestire i volumi virtuali.

I volumi risultano quindi completamente indipendenti sia dai server sia dai sottosistemi storage sulla SAN.

Tutto lo storage fisico viene reso visibile ed è disponibile per la creazione e allocazione di volumi virtuali.

4.2 I modelli di virtualizzazione in-band e out-of-band

Nel modello *in-band* i dati e il loro flusso di controllo utilizzano lo stesso percorso; tale modello può essere considerato più facile da realizzare poiché non richiede l'installazione di software specifico sui server.

La virtualizzazione *in-band* rende disponibili soluzioni avanzate e di caching dei dati all'interno della rete storage e scarica l'host dallo svolgimento di alcune funzioni, consentendo inoltre di supportare più host eterogenei e di prevedere l'ottimizzazione delle prestazioni in relazione al percorso dei dati.

Per realizzare una soluzione *out-of-band*, invece, è richiesto un server di controllo aggiuntivo e il percorso dei dati viene separato dal flusso di controllo; anche i dati sono divisi dai metadati.

Anche questo sistema supporta host eterogenei, permettendo così all'azienda di poter scegliere fra più fornitori storage.

Questo metodo, inoltre, offre un'ottima scalabilità ed una buona integrazione con il software per la gestione dello storage.

Aggiungendo ulteriori componenti di virtualizzazione è però necessario tener conto di un aumento delle complessità e del livello di affidabilità.

I due modelli di virtualizzazione possono anche coesistere; la scelta del sistema più adatto da adottare dipende dalle specifiche esigenze aziendali.

L'aspetto assolutamente positivo è che *le tecniche di virtualizzazione hanno raggiunto un elevato livello di maturità*; in molte aziende esse sono già utilizzate o comunque sono tra le prime priorità di chi ha la responsabilità dei sistemi informativi.

4.3 Partizionamento

L'utilizzo di un unico pool di risorse anziché di singoli sistemi porta benefici, ma anche criticità; pensiamo, infatti, alla sicurezza nel gestire dati differenti, ossia di diversa natura: online, statici, dinamici e legati a soggetti diversi, originati o trattati da differenti applicativi residenti in risorse consolidate non più separate.

Per affrontare questa criticità, in molte soluzioni di virtualizzazione è presente la funzionalità di partizionamento, che, conservando intatta la possibilità di riunire tutte le risorse storage in un unico pool, lascia la possibilità di separare i diversi ambienti e le diverse tipologie di informazioni. Si possono così creare dei "partizionamenti" di risorse ognuna dedicato ad uno scopo diverso che può essere il singolo cliente da servire, il singolo servizio in outsourcing da erogare, il singolo applicativo da eseguire, introducendo così un ulteriore livello di sicurezza.

5. Tiered Storage

Il concetto di Tiered Storage è spesso usato come sinonimo di *Hierarchical Storage Management* (HSM); le differenze sono però sostanziali.

Gran parte dello storage disponibile risulta occupato da dati consultati raramente.

Per garantire il più favorevole rapporto tra costi ed esigenze (velocità) di accesso ai dati, buona parte di questi possono essere posizionanti su dispositivi più economici e con prestazioni inferiori, mentre i dati più richiesti sono memorizzati su dispositivi più rapidi: è questo il concetto del HSM, cioè della gestione gerarchica dello storage.

Per raggiungere i suoi obiettivi, in modo trasparente per l'utente o per l'applicativo, un sistema di HSM sposta i dati, in funzione della loro frequenza di accesso, dall'area storage di primo livello a quella di secondo o terzo livello con l'obiettivo di rendere minima l'occupazione dello storage a più alte prestazioni e quindi contenere i costi.

Un primo livello è in genere rappresentato da soluzioni a disco, il secondo e terzo livello possono essere costituiti, ad esempio, da dischi di tecnologia a più basse prestazioni o nastri, il terzo livello da librerie nastro; questo tipo di gestione differenziata permette di sfruttare le prestazioni dei sistemi a disco e i bassi costi e consumi energetici dei sistemi basati su nastro.

Esistono più soluzioni HSM che si caratterizzano per come viene gestito l'automatismo e per il suo grado di trasparenza.

Spostare i dati da un supporto ad un altro solo in funzione di quanto spesso questi sono utilizzati crea problemi nei casi in cui le aziende necessitano di accedere immediatamente ad informazioni che normalmente sono poco richieste.

Una metodologia utilizzata per risolvere questo problema è quella che prevede di catalogare i dati/informazioni in gruppi sulla base

dei parametri che determinino il valore per l'azienda²⁵ ("valore di business") e di individuare i livelli (*Tier*) di storage da adottare (in genere da 2 ad un massimo 4) e con le loro caratteristiche in termini di prestazioni, funzioni e costi; si potranno quindi stabilire le regole per attribuire i gruppi di informazioni ai vari livelli di storage in modo da semplificare la gestione dello storage e ridurre sia i costi sia i rischi associati all'errore umano.

I livelli di storage non vengono individuati in base a caratteristiche strutturali/architetturali o in funzione del fornitore, ma esclusivamente in funzione delle prestazioni, della capacità, delle funzionalità di cui dispongono e del prezzo.

L'obiettivo principale del *Tiered Storage* è, quindi, quello di ridurre i costi globali dello storage, ottimizzando i tempi d'accesso ai dati rilevanti per il business pur mantenendone la corretta conservazione per il tempo dovuto o voluto.

Assegnare i dati ad un particolare tipo di storage può essere (il più delle volte lo è) un'attività sufficientemente complessa; per questo motivo alcuni fornitori offrono software in grado di gestire tale processo sulla base di definite politiche aziendali.

In conclusione una considerazione, valida in particolare per le aziende di minori dimensioni che normalmente non hanno ambienti storage estremamente eterogenei in termini di produttori o tecnologie, ma sicuramente devono gestire acquisti dilazionati nel tempo e quindi l'uso contemporaneo di tecnologie nuove e più datate.

²⁵ Parametri per determinare il valore di un dato/informazione possono essere la sua criticità (mission critical, business critical, ...), il tipo (transazionale, non strutturato, e-mail, immagine, ...), il settore cui fa riferimento (marketing, vendite, produzione, personale, amministrazione e finanza, ricerca e sviluppo, ...) o l'anzianità della stessa, il livello di protezione richiesto, la riservatezza attribuita, il tempo minimo di accesso richiesto, le esigenze di disponibilità, i volumi ed il loro incremento annuo, ...

Le tecniche di *Tiered Storage* offrono l'opportunità di sfruttare appieno gli investimenti fatti, utilizzando, per esempio, la porzione di risorse formalmente obsolete per svolgere il ruolo di secondo livello molto utile a conservare le informazioni nel lungo periodo.

6. Evoluzioni future

I miglioramenti di capacità e prestazione delle memorie digitali sono stati ottenuti aumentando continuamente la densità con cui venivano memorizzate le informazioni sui supporti. Questo processo di miglioramento continuo è multi-disciplinare e coinvolge diverse aree di ricerca dalla scienza dei materiali, dalle teorie sui codici, dall'elettromagnetismo, alla microelettronica.

Nel 2006, ad esempio, nei laboratori di ricerca sono stati dimostrati prototipi di nastro magnetico che potrebbero consentire la realizzazione nei prossimi anni di cartucce da 8 TB e sono attivi progetti di ricerca per raggiungere in futuro capacità di 25 TB per cartuccia.

È difficile stimare il futuro delle tecnologie magnetiche perché ovviamente, prima o poi, verrà raggiunto il limite fisico oltre il quale non sarà più possibile migliorare ulteriormente la densità di memorizzazione. Tuttavia, negli anni passati, ogni volta che veniva raggiunto un limite fisico rapidamente la ricerca ha trovato un'innovazione tecnologica in grado di superare tale limite.

In ogni caso nei laboratori di ricerca universitari e delle principali aziende di tecnologia sono attivi da anni diversi progetti di ricerca e sviluppo per tecnologie di memoria alternative con le caratteristiche di capacità, prestazioni e affidabilità richieste.

A solo titolo di esempio di seguito sono alcuni dei progetti più

interessanti che sono stati presentati negli ultimi anni:

- Memorie *olografiche*: utilizzo di un elemento di memoria cristallino tridimensionale che viene scritto e letto sfruttando l'interazione tra fasci di luce polarizzata. A tal proposito si ricorda che già nel 1997 si considerava questa tecnologia pronta per essere commercializzata nel giro di pochi anni, ma solo con il 2006 sono comparse le prime realizzazioni concrete, a livello però di prototipi o poco più.
- Memorie *Magnetic RAM (MRAM)*: sviluppo di chip di memoria contenenti una matrice di micro-dispositivi magnetici²⁶.
- Memorie *nanomeccaniche*: dispositivi nano-tecnologici in grado di modificare fisicamente il supporto di registrazione mediante nano-incisioni²⁷ o di apportare modifiche della struttura molecolare, se non a quella atomica, del materiale²⁸.

In sintesi, quindi, *l'evoluzione delle singole tecnologie e la possibile nascita di tecnologie alternative con capacità e costi migliorati, deve essere considerata in un progetto di archiviazione di lungo periodo in quanto il ciclo di vita delle informazioni può superare quello delle tecnologie.*

7. Alcuni esempi di realizzazioni RAID

7.1 RAID 4

Questa tipologia di RAID prevede la presenza di un disco di parità orizzontale per ogni gruppo RAID, in tal modo a fronte

del malfunzionamento di una unità disco, non vengono persi i dati del volume di cui il disco fa parte. In caso di malfunzionamento del disco, pertanto, il sistema continua a funzionare con i dati intatti e senza nessuna perdita in termini di prestazioni.

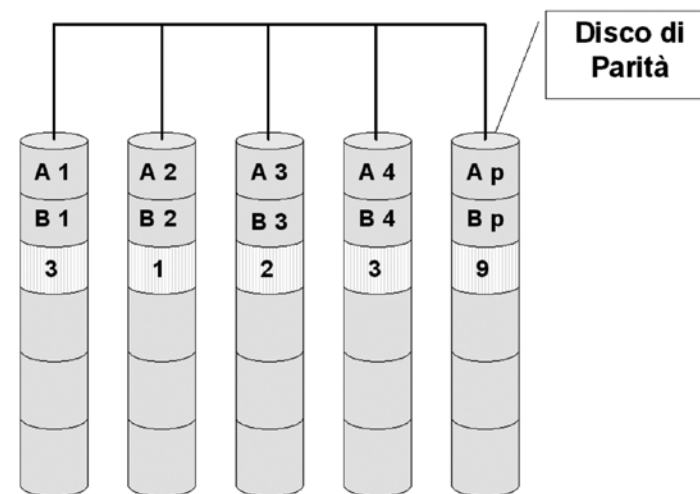


Figura 5 – Schema di RAID 4

La figura 5 mostra un RAID 4 tradizionale con l'ultimo disco come disco di parità. La terza riga nella figura è stata popolata con dei dati d'esempio ed il valore di parità calcolato dai dati è stato registrato sul relativo settore del disco di parità.

In questo caso, la modalità con cui la parità è stata calcolata consiste nel sommare i valori dei singoli settori della riga, quindi salvare il risultato come valore di parità ($3 + 1 + 2 + 3 = 9$)²⁹.

Se si presentasse la necessità di ricostruire il dato per il danneggiamento di un singolo disco, il processo utilizzato per la crea-

²⁶ www.research.ibm.com/resources/news/20001207_mramimages.shtml.

²⁷ domino.watson.ibm.com/comm/pr.nsf/pages/news.20020611_millipede.html.

²⁸ Atomic level storage: www.ibm.com/press/us/en/pressrelease/22254.wss.

²⁹ Nella pratica la parità è calcolata tramite il costrutto OR (XOR), ma la somma è un processo abbastanza simile, e funziona bene ai fini di questo esempio.

zione della parità sarebbe semplicemente invertito; ad esempio, se si dovesse guastare il primo disco, il RAID 4 per ricostruire il valore 3 nella prima colonna della figura sottrarrebbe i valori dei dischi rimasti ($9 - 3 - 2 - 1 = 3$). Questo esempio di ricostruzione tramite singola parità, dovrebbe aiutare a comprendere il concetto del perché un dato è protetto, ma non dal guasto di più di un disco.

7.2 RAID 5

Un sistema RAID 5 usa una divisione dei dati a livello di blocco (Ax, Bx, Cx, Dx, ...) con i dati di parità (Ap, Bp, Cp, Dp, ...) distribuiti tra tutti i dischi appartenenti al RAID. Questa è una delle realizzazioni più diffuse. Virtualmente ogni sistema di storage prevede il RAID-5 tra le sue opzioni.

Nell'esempio in figura 6, una richiesta al blocco A1 potrebbe essere evasa dal disco 1.

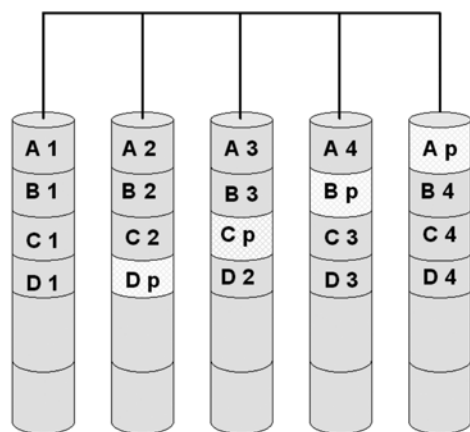


Figura 6 – Schema di RAID 5; i dati di parità sono distribuiti su più dischi

Una simultanea richiesta per il blocco B1 dovrebbe aspettare, ma una richiesta simultanea per il blocco B2 potrebbe essere evasa in contemporanea.

Ogni volta che un blocco di dati deve essere scritto nel sistema di dischi, un blocco di

parità è generato all'interno della stripe; se un altro blocco, o qualche porzione dello stesso blocco, è scritta nella stessa stripe, il blocco di parità viene ricalcolato e riscritto.

Il disco usato per memorizzare le parità viene modificato tra una stripe e la successiva; in questo modo si riescono a distribuire i blocchi di parità.

In un sistema RAID 5 che ha un solo blocco di parità per stripe, la rottura di un secondo disco comporta, però, la perdita di tutti i dati presenti nel sistema.

Nelle realizzazioni con più di 14 dischi, o in situazioni dove è necessaria un'elevata ridondanza dei dati, viene usata spesso una forma di RAID 5 con doppia parità (detta anche RAID 6), che riesce a gestire il guasto contemporaneo di due dischi.

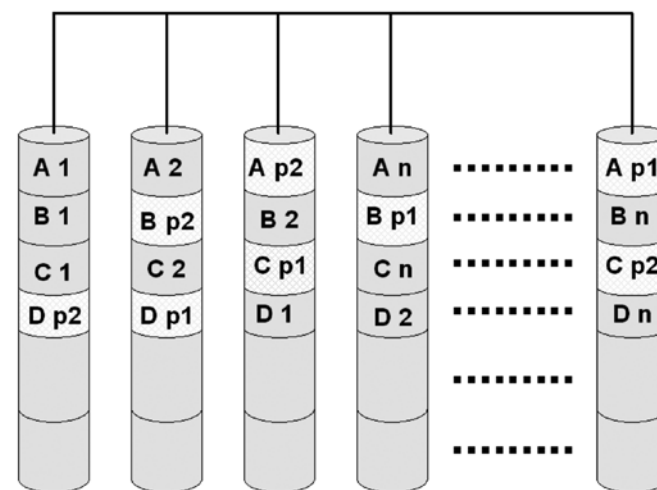


Figura 7 – Schema di RAID 6; è una forma di RAID 5, ma con calcolo e registrazione di un doppio blocco di parità.

Capitolo V

Chi può conservare la memoria digitale?

1. Introduzione

Il problema della conservazione della memoria digitale è certamente molto complesso e non si ha a disposizione (se mai la si avrà) una soluzione definitiva e permanente che ci consenta di realizzare soluzioni con un grado di affidabilità e di certezza degli strumenti da adottare come ci piacerebbe.

Per non correre il rischio di perdere la nostra memoria l'unica via da percorrere, oggi, è quella di adottare le tecniche e le tecnologie di cui disponiamo che ci consentono di proteggere e conservare il patrimonio digitale che ci interessa, ben sapendo che in un futuro, anche abbastanza prossimo, si dovrà quantomeno trasferire le nostre informazioni digitali su supporti basati su nuove e più complesse tecnologie, proteggendo non solo l'informazione in sé, ma anche le correlate informazioni che ne garantiscono l'autenticità, l'origine e ne descrivono il contesto.

2. Il problema

Pur nei limiti caratteristici di un momento di transizione, le grandi aziende, organizzazioni ed amministrazioni si sono già attrezzate per garantirsi la protezione della memoria digitale.

Le piccole sono più in difficoltà, se non proprio assolutamente impreparate e, per conservare i loro dati in formato digitale, si stanno sempre più affidando a *servizi esterni*.

Questa che è oggi una tendenza potrebbe diventare la regola.

Attualmente queste aziende hanno ben poca protezione nei confronti di fornitori non opportunamente qualificati (si ricorda che questi fornitori hanno il delicato compito di conservare la memoria delle aziende con tutte le conseguenze che derivano in caso di perdita di tutta o parte della stessa; non solo, ma anche di offrire una ragionevole garanzia di essere in grado di mantenere l'impegno per tutto il tempo in cui l'informazione deve essere conservata).

Premesso che la piccola utenza non ha le competenze e la capacità (forse non è neanche corretto pretendere che le abbia) per valutare l'affidabilità di un fornitore di questo tipo di servizi, il solo costo (strettamente legato alla qualità del servizio offerto) potrebbe rappresentare il parametro di scelta principale inducendo a preferire fornitori di qualità più bassa rispetto a quelli in grado di offrire maggiore qualità e pertanto maggiori garanzie, ma anche maggior costo.

Nella pratica di tutti i giorni si può poi facilmente riscontrare che un sempre crescente numero di società si stanno proponendo come prestatori di tale tipologia di servizio, in particolare con riferimento alle attività di Conservazione Sostitutiva di documenti rilevanti ai fini tributari e di fatturazione elettronica.

Purtroppo anche da un'analisi superficiale si ricava l'impressione che alcuni di questi fornitori non abbiano compreso non solo la sostanza delle norme che regolano le succitate attività (basta vedere la terminologia adottata quale, ad esempio, "archiviazione ottica" a volte integrata con l'attributo "sostitutiva"), ma anche le elementari norme di sicurezza che esse sottendono. Senza, poi, approfondire la comunicazione di marketing a volte pericolosamente fuorviante.

I manuali della conservazione, scendendo poi in argomenti più tecnici, contengono di tutto e di più: dall'enfatizzazione del prodotto adottato ai moduli per registrare le più svariate attività, ma che non indicano se il processo offerto è in sintonia con misure di sicurezza informatica diffusamente riconosciute quali lo standard ISO/IEC 27001, l'ISO/IEC 17799 o il precedente BS/7799.

Vi sarà quindi il rischio che gli utenti di fornitori di servizi di qualità inferiore possano avere la spiacevole sorpresa, nell'arco di alcuni anni, di scoprire di non essere in grado di esibire alcuni documenti conservati in maniera sostitutiva, in particolare quelli fiscali, o che essi non rispecchino le norme relative.

Questo scenario è certamente da rifiutare.

Una certa linea di pensiero sostiene che il mercato, con quelli che si ritiene possano essere i suoi propri meccanismi di autodifesa, riesca a selezionare chi fornisce questi servizi in modo affidabile, per cui a regime solo i migliori sopravviverebbero e continuerebbero ad operare.

Potrebbe anche essere vero, ma vale la pena di far correre questo rischio alle aziende di minor dimensione che sono, peraltro, la struttura portante dell'impresa nazionale?

Qui viene in supporto l'Unione Europea, molto attenta a tutti gli aspetti di equilibrio dei mercati e di protezione dell'utenza, con la direttiva 2006/123/CE del Parlamento Europeo e del Consiglio del 12 dicembre 2006 relativa ai servizi nel mercato interno che nella premessa (97)³⁰ esplicitamente dichiara l'esigenza di

³⁰ Occorre prevedere nella presente direttiva delle norme relative all'alta qualità dei servizi, che soddisfino in particolare requisiti di informazione e trasparenza. Tali norme dovrebbero applicarsi sia nel caso di prestazioni di servizi transfrontalieri tra Stati membri, sia nel caso di servizi offerti da un prestatore all'interno dello Stato membro in cui egli è stabilito senza imporre inutili

regole che favoriscano i servizi di più elevata qualità favorendo informazione e trasparenza, per poi all'articolo 26³¹ esplicitare l'obbligo, per i Paesi, di incoraggiare i prestatori a garantire su base volontaria la qualità dei servizi, indicando gli strumenti per definirli e farli conoscere.

3. Una soluzione

È pertanto auspicabile che venga introdotto, almeno inizialmente, per i fornitori di servizi, oggi detti di Conservazione Sostitutiva, un *sistema di accreditamento facoltativo*; ciò consentirebbe alle piccole aziende di individuare fornitori a cui sia stata riconosciuta, da un ente terzo ed affidabile, la capacità di prestare servizi aderenti a ben precisi livelli di prestazioni.

oneri alle piccole e medie imprese. Esse non dovrebbero impedire in nessun caso agli Stati membri di applicare, conformemente alla presente direttiva e ad altre norme comunitarie, requisiti di qualità supplementari o diversi.

³¹ Articolo 26 - Politica in materia di qualità dei servizi

1. Gli Stati membri, in collaborazione con la Commissione, adottano misure di accompagnamento volte ad incoraggiare i prestatori a garantire, su base volontaria, la qualità dei servizi, in particolare:

a) facendo certificare o valutare le loro attività da organismi indipendenti o accreditati;
b) elaborando una carta di qualità propria o aderendo alle carte o ai marchi di qualità messi a punto da organismi e ordini professionali a livello comunitario.

2. Gli Stati membri provvedono affinché le informazioni sul significato di taluni marchi e sui criteri di attribuzione dei marchi e di altri attestati di qualità relativi ai servizi siano facilmente accessibili ai prestatori e ai destinatari.

3. Gli Stati membri, in collaborazione con la Commissione, adottano misure di accompagnamento volte ad incoraggiare gli ordini professionali, le camere di commercio e artigianato e le associazioni dei consumatori negli Stati membri a collaborare a livello comunitario per promuovere la qualità dei servizi, in particolare facilitando il riconoscimento della qualità dei prestatori.

4. Gli Stati membri, in collaborazione con la Commissione, adottano misure di accompagnamento volte ad incoraggiare lo sviluppo della comunicazione critica, in particolare da parte delle associazioni dei consumatori, relativa alle qualità e ai difetti dei servizi, in particolare lo sviluppo a livello comunitario di prove o collaudi comparativi e della comunicazione dei loro risultati.

5. Gli Stati membri, in collaborazione con la Commissione, incoraggiano lo sviluppo di norme volontarie europee intese ad agevolare la compatibilità fra servizi forniti da prestatori di Stati membri diversi, l'informazione del destinatario e la qualità dei servizi.

La volontarietà di questo meccanismo non escluderebbe alcun fornitore dal mercato, ma darebbe anche all'utenza sprovvista della competenza specifica la possibilità di individuare e riconoscere i fornitori più qualificati.

Questo meccanismo di accreditamento si fa sempre più importante e necessario anche per le aziende di maggiori dimensioni, specialmente oggi che ci sono tutte le condizioni per uno sviluppo dei processi di fatturazione elettronica che, tra l'altro, hanno come conseguenza diretta un'attività di Conservazione Sostitutiva.

Non a caso la legge finanziaria 2008³², che prevede la fatturazione elettronica per tutte le forniture a favore delle amministrazioni dello Stato e degli enti pubblici nazionali, all'articolo 213, comma 1, punto e), prevede esplicitamente l'impiego di "intermediari *abilitati*" per supportare le aziende e le amministrazioni che lo necessitano, rispettivamente, nella generazione/invio e nella ricezione/acquisizione delle fatture elettroniche.

Attualmente nel nostro Paese la conservazione a norma della memoria digitale viene realizzata con l'impegno della tecnica della firma digitale, secondo quanto prescritto per la cosiddetta Conservazione Sostitutiva; cosa questa che permarrà con buona probabilità anche nel futuro.

Anche trascurando la citata direttiva 2006/123/CE, si può concludere che un meccanismo di accreditamento volontario per la conservazione della memoria digitale poteva già essere istituito nel nostro Paese se se ne fosse compresa l'esigenza.

³² Legge 24 dicembre 2007 n. 244, articolo 1, commi dal 209 al 214.

Già la direttiva 1999/93/CE alla premessa (9) recitava: “le firme elettroniche verranno usate in svariate circostanze ed applicazioni, che comporteranno un’ampia gamma di nuovi servizi e prodotti facenti uso di firme elettroniche o ad esse collegati; la definizione di tali prodotti e servizi non dovrebbe essere limitata al rilascio e alla gestione di certificati, ma *comprenderebbe anche ogni altro servizio e prodotto facente uso di firme elettroniche, o ad esse ausiliario*, quali servizi di immatricolazione, servizi di apposizione del giorno e dell’ora, servizi di repertorizzazione, servizi informatici o di consulenza relativi alle firme elettroniche”.

La medesima direttiva all’art. 2 comma 11, poi, definisce il “prestatore di servizi di certificazione” come un’entità o una persona fisica o giuridica che rilascia certificati o “fornisce altri servizi connessi alle firme elettroniche”.

Il comma 2 dell’articolo 3 della direttiva, infine, afferma che: “gli Stati membri possono introdurre o conservare sistemi di accreditamento facoltativi volti a fornire servizi di certificazione di livello più elevato.”.

Per quanto concerne la normativa italiana non si fa esplicitamente riferimento ad un prestatore di servizi di certificazione; il decreto legislativo 82/2005 (Codice dell’amministrazione digitale) all’art. 1 comma 1, lettera g) definisce, però, il certificatore come il soggetto che presta servizi di certificazione delle firme elettroniche o “che fornisce altri servizi connessi con queste ultime”.

Non tragga, dunque, in inganno il termine “certificatore” che, impiegato forse per la sua maggior brevità rispetto al termine “prestatore di servizi di certificazione”, pur stabilendo un legame “istintivo” con la sola funzione di “rilascio di certificati digitali”, include anche prestatori di servizi non collegati al

rilascio di certificati purché siano “servizi connessi alle firme elettroniche”.

Quanto sopra esposto, anche se non in modo del tutto esplicito, non esclude, anzi fa ragionevolmente ritenere che anche in Italia potevano già essere introdotti sistemi di accreditamento facoltativo per prestatori di servizi connessi alle firme elettroniche che non consistano nel mero rilascio di certificati.

4. Alcuni suggerimenti

Di seguito, a titolo esemplificativo, è un elenco non esaustivo dei requisiti che potrebbero essere alla base dell’accreditamento di un fornitore di servizi di conservazione della memoria digitale:

- requisiti societari sia dal punto di vista di solidità economica sia dal punto di affidabilità degli amministratori;
- requisiti tecnici, basati su esperienza concreta, in materia di sicurezza, gestione di un data center, gestione di programmi di business continuity e di disaster recovery, archivistica, internal audit, ecc.;
- disponibilità di una struttura con identificati i responsabili del servizio nel suo complesso, dei servizi tecnici, della sicurezza, dell’audit interno, della business continuity o almeno del disaster recovery, dell’archivistica ove necessario, ...

È, inoltre, necessario predisporre:

- un piano per la sicurezza, basato su un affidabile *Information Security Management System* che tragga origine da un *Risk Assessment* esibibile e sostenibile;
- sistemi di protezione adeguati ai dati conservati e che dispongano delle caratteristiche necessarie a proteggere da attacchi gli ambienti di trasferimento dei documenti sui

supporti per la conservazione, gli ambienti di conservazione nel lungo periodo e quelli di verifica delle firme (firewall, antivirus, antispymware, ... aggiornati con continuità);

- un piano di business continuity e di disaster recovery;
- un manuale del responsabile tecnico, operativo, ecc.;
- un manuale in cui siano descritti ben chiariti
 - a) le attività e le misure, tecniche e di conformità alle norme vigenti, a carico del fornitore di servizi per assicurare la conservazione delle informazioni e per quanto tempo;
 - b) i termini della responsabilità del fornitore di servizi e del cliente.

È infine indispensabile provvedere ad un'assicurazione contro la perdita di dati.

Dal momento infine che la conservazione della memoria digitale può prevedere tempi a volte più lunghi di una ragionevole vita aziendale, il sistema di accreditamento dovrebbe prevedere che, al verificarsi per un qualunque motivo della cessazione di attività di una azienda, il patrimonio informativo digitale da conservare passi in carico ad una o più aziende che hanno ottenuto lo stesso accreditamento e, quindi, in grado di garantire uno stesso livello qualitativo del servizio.

Capitolo VI

Sistemi di Conservazione Sostitutiva: tecnologia e funzionalità

di Andrea Venturato³³

1. Gli archivi digitali

1.1 Documenti digitali e archivi digitali

Grazie alla costante evoluzione informatica, enti e aziende hanno assistito a una crescita esponenziale del numero di documenti prodotti, ricevuti e di conseguenza archiviati.

Con la diffusione dell'Office Automation e dei sistemi di stampa personali o dipartimentali, è infatti diventato estremamente facile e accessibile a tutti redigere e stampare documenti.

È nato così un nuovo tipo di documento: il documento digitale, che ha aperto nuove opportunità in termini di efficienza operativa dei processi di business correlati e di semplicità di fruizione, condivisione e distribuzione dell'informazione e della conoscenza.

I documenti digitali hanno portato con sé lo sviluppo dei sistemi di archiviazione elettronica prima e di gestione documentale

³³ Andrea Venturato, laurea in Ingegneria Elettronica presso l'Università degli Studi di Padova, da oltre 10 anni si occupa dello sviluppo di soluzioni innovative di gestione elettronica documentale in ambito pubblico e privato. Attualmente riveste il ruolo di Innovation & Competence Center Executive presso Siav S.p.A.

poi: oggi ai tradizionali archivi cartacei si affiancano, ben più ricchi di contenuti informativi e largamente condivisi, gli archivi digitali, in cui confluiscono anche forme documentali proprie dell'Information Technology quali ad esempio i messaggi di posta elettronica, le pagine web e i filmati.

Sulla spinta di tali fenomeni, sono state emanate negli ultimi anni una serie di norme in ambito europeo e nazionale, con l'obiettivo di dare valore legale al documento digitale in sede giudiziaria e renderlo opponibile a terzi.

Il documento digitale diventa così sostitutivo, a tutti gli effetti, del documento cartaceo.

Concentrandosi sugli aspetti tecnici e procedurali, la normativa ad oggi sembra non prendere sufficientemente in considerazione gli aspetti – pur strategici per le aziende - connessi alla *fruibilità* dei documenti digitali e alla *gestione* dell'insieme dei documenti sostitutivi.

1.2 Gestire la complessità: dallo strumento al sistema

Per dare risposta alla normativa si sono sviluppate soluzioni in grado di supportare il processo di conservazione, tipicamente attraverso strumenti che gestiscono la fase di produzione dei supporti fisici a norma di legge.

I documenti digitali entrano nel processo, vengono sottoposti alle verifiche formali, vengono in alcuni casi firmati digitalmente e vanno a costituire dei supporti certificati: nella maggior parte dei casi tali supporti vengono riposti in *caveau* e ne escono, e con essi i documenti, solo in caso di accertamento da parte delle autorità o per le verifiche periodiche di validità. Nel frattempo però le organizzazioni continuano a lavorare con copie non so-

stitutive dei documenti o, peggio ancora, con copie cartacee.

La normativa e la sua applicazione sono un fatto ancora recente, ma chi gestirà nei prossimi anni il contenuto dei *caveau*?

Come si farà a sapere con esattezza cosa contengono, cosa si può scartare o cosa sta diventando obsoleto?

A differenza di un archivio cartaceo infatti, un archivio digitale è composto di dati informatici e questo apre un nuovo scenario applicativo, quello dei *sistemi per la gestione* dell'insieme dei documenti sostitutivi.

Questi sistemi non sono semplicemente strumenti in grado di rendere legali i documenti digitali o elettronici ma costituiscono veri e propri *custodi* degli archivi digitali, garantendo negli anni la loro integrità e, soprattutto, rendendo fruibili le informazioni conservate.

1.3 I rischi degli archivi digitali

Il documento cartaceo si può definire come un oggetto fisico che consente di accedere direttamente all'informazione in esso contenuta ricorrendo esclusivamente all'intelletto e al proprio bagaglio di conoscenza.

Il documento digitale è invece una rappresentazione astratta e numerica della realtà sensoriale: l'informazione è codificata, memorizzata su supporti fisici discontinui e, di fatto, non direttamente accessibile.

Ancora più indirettamente accessibili sono le componenti di rappresentazione dell'informazione, (come i font, la formattazione, l'impaginazione, ...) attraverso cui la stessa deve essere "traddotta" con l'ausilio di uno strumento informatico (hardware e

software) per risultare intelligibile.

Negli ultimi decenni, l'informatica si è sempre più perfezionata in questo, migliorando la facilità e flessibilità nella definizione, nella produzione e nel riuso.

I documenti digitali, e di conseguenza gli archivi che li contengono, sono soggetti a potenziali rischi o minacce che riguardano sia la forma memorizzata sia la forma rappresentata.

Le minacce possono riguardare:

- *l'esistenza*: quando il documento digitale non è più disponibile, di fatto distrutto.
- *l'alterazione*: quando il documento digitale viene modificato nella sua forma memorizzata o nella sua forma rappresentata; l'informazione non è più fedele all'originale.
- *la fruibilità*: quando il documento digitale esiste ma non è di fatto utilizzabile. Si tratta dunque dell'impossibilità di rappresentarlo.

Come è facile intuire, gli archivi cartacei sono soggetti ai medesimi rischi: si pensi ad incendi, allagamenti o smarrimento (esistenza), a truffe (alterazione), alla difficoltà di accedervi e rintracciarli (fruibilità).

Nonostante lo scetticismo di molti è più facile proteggere gli archivi digitali che gli archivi cartacei: questo è dovuto principalmente al loro ridottissimo volume fisico e al ridotto costo di memorizzazione.

2. Sistemi per la Conservazione Sostitutiva: requisiti e funzionalità di base

Le funzioni più critiche del processo di Conservazione Sostitutiva sono legate alla validità del documento, specie se tale docu-

mento è portatore di atti potenzialmente rilevanti.

Di seguito vengono descritti i requisiti e le funzionalità generiche di un sistema per la Conservazione Sostitutiva, in cui i sottosistemi di verifica, consolidamento, conservazione sono sostanzialmente visti come funzioni integrate.

I documenti firmati vengono affidati al sistema da parte delle applicazioni che li producono, quindi proseguono il proprio ciclo sino alla generazione di distinte sequenze di supporti virtuali che vengono immagazzinate per successive e differenti procedure di certificazione.

Il sistema, a livello centralizzato, deve prendere in carico i documenti informatici firmati digitalmente nell'ambito delle varie applicazioni ed eseguire:

- verifica accurata delle firme digitali apposte al documento;
- consolidamento del valore probatorio;
- conservazione su supporti a norme CNIPA (delib. 11/2004).

Per rispondere a queste specifiche, il sistema deve eseguire per ogni sottoscrizione elettronica eventualmente presente sul documento, le seguenti operazioni (fase di consolidamento):

- verifica della conformità alle normative e regole tecniche vigenti (DLGS 82/2005 "CAD" e successive modificazioni; DPCM 13/01/2004; AIPA/CR 24/2000) del documento inoltrato, in termini di formato, *naming convention*, contenuti della envelope, certificati, ecc.
- verifica dell'integrità del documento: il documento non può essere stato alterato dopo l'apposizione della firma;
- verifica che il certificato di chiave pubblica di ogni sottoscrittore (allegato obbligatoriamente al documento) sia effettivamente un certificato qualificato emesso da uno qualsiasi dei certificatori accreditati dal CNIPA. Tale verifica viene sod-

disfatta utilizzando come *trust point* la lista dei certificati delle chiavi di certificazione pubblicata dal CNIPA e firmata dal suo presidente; il sistema di verifica deve prevedere un aggiornamento automatico di questo elenco con periodicità configurabile da 1 a 30 giorni.

- attivazione di funzioni di sicurezza per proteggere l'integrità del trust point e verifica – con riferimento anche al relativo stato di revoca – la firma del presidente del CNIPA, usando i riferimenti pubblicati nella circolare CNIPA/CR 42/2003, ad ogni avvio.
- gestione dell'eventuale cambiamento del vertice CNIPA (quindi possibile revoca del certificato e nuovo firmatario dell'elenco pubblico) tramite aggiornamento del software e consentendo l'operatività nel frattempo (situazione definita “warning”);
- verifica che il certificato di chiave pubblica di ogni sottoscrittore non sia scaduto;
- verifica che il certificato di chiave pubblica di ogni sottoscrittore non sia in stato di sospensione o revoca. Questa situazione è particolarmente delicata, poiché il tempo intercorrente tra una richiesta di revoca e l'effettiva visibilità della sospensione o revoca nella relativa CRL non è affatto nullo. Il sistema deve rispettare un *grace period* configurabile, attendendo la pubblicazione di eventuali nuove edizioni della CRL di riferimento prima di considerare una firma certamente valida.
- controllo dello stato di revoca di un certificato prelevando la CRL dal CRL Distribution Point contenuto nel certificato stesso. Qualora la lettura di una CRL sia temporaneamente impossibile a causa di un problema di indisponibilità (servizio di rete, server della CA) il sistema deve gestire l'attesa e risolvere il problema al superamento del transitorio, emettendo opportune notifiche in caso di impossibilità perdurante.

In caso l'iter di verifica riveli il fallimento di uno dei passi descritti, il sistema di consolidamento deve prevedere l'invio di un messaggio di notifica al responsabile della conservazione.

Per poter gestire la conservazione a lungo termine dei documenti informatici, in accordo con la normativa sulla Conservazione Sostitutiva (deliberazione 11/2004 CNIPA), il sistema deve disporre delle seguenti caratteristiche funzionali (fase di conservazione):

- supporto di volumi di archiviazione “di lavoro” su cui si accumulano documenti fino al raggiungimento di una soglia (dimensionale o temporale) per la chiusura del singolo volume (processo di conservazione, delibera 11/2004 art. 3);
- metadati di classificazione definibili dall'utente;
- gestione automatica del consolidamento dei volumi che hanno raggiunto la dimensione o l'età programmate per la conservazione, con relativa creazione di nuovi volumi vuoti per l'accumulo di successivi documenti;
- notifica via e-mail al responsabile della conservazione della disponibilità di ogni nuovo volume congelato, per il suo intervento in fase di chiusura del volume (firma digitale, marca temporale);
- interfaccia utente per la navigazione attraverso i volumi, la visualizzazione dei contenuti, ricerche, esportazioni, stampa.
- opzioni che consentono di scegliere, come supporti di materializzazione, i diversi media ottici. In prima approssimazione è possibile operare con CD-R e DVD+R/DVD-R, anche a doppio strato;
- possibilità di gestire le operazioni di materializzazione in modalità manuale (on demand) oppure automatica mediante l'utilizzo di hardware specialistico (robot di masterizzazione) per la materializzazione massiva, compresa la produzio-

ne di copie di sicurezza etichettate;

- al momento del congelamento in vista della chiusura, possibilità di esportare i documenti su porzioni di file system, in modo da rendere i volumi di conservazione ispezionabili anche senza la disponibilità di speciali strumenti di gestione database;
- il file che rappresenta l'evidenza informatica contenente le impronte, come definito dalla deliberazione CNIPA 11/2004 art. 3, è un documento informatico firmato e dal contenuto leggibile con strumenti standard (formato XML);
- compatibilità con altri tipi di riferimento temporale "esterno", oltre a quello di default (marcatura temporale);
- i documenti contenuti nel volume sono accessibili ad applicazioni esterne mediante l'utilizzo di web service di integrazione;
- possibilità di memorizzare i metadati sia nel file firmato contenente anche le impronte, sia in file di database apribili con strumenti di uso comune;
- possibilità di gestire la firma del pubblico ufficiale (da remoto) in aggiunta a quella del responsabile della conservazione, in conformità all'art. 4 della deliberazione 11/2004.

2.1 Obblighi del responsabile della conservazione

Sulla base alle raccomandazioni CNIPA presenti nell'Art. 5 della delibera 11/2004 il responsabile della conservazione è tenuto a svolgere i seguenti compiti:

- definire, realizzare, documentare ed aggiornare periodicamente (comunque ad ogni variazione) le procedure di archiviazione, di protezione dati e di protezione da eventi disastrosi;

- verificare periodicamente la corretta esecuzione delle procedure di archiviazione mediante controlli a campione sulla validità dei dati registrati;
- verificare periodicamente mediante prove (almeno con cadenza annuale) la capacità di ripristino dei dati coinvolgendo il personale addetto;
- documentare l'esito delle prove e delle eventuali azioni per la correzione e/o miglioramento del processo;
- verificare periodicamente (almeno ogni 3 anni) l'obsolescenza tecnologica dei dispositivi utilizzati, del formato e del software di accesso ai dati al fine di tempificare opportunamente le attività di riversamento.

La soluzione di Conservazione Sostitutiva deve consentire di automatizzare il più possibile le verifiche richieste, producendo la necessaria documentazione a corredo.

2.2 Consolidamento del valore probatorio: approfondimento

Al documento firmato digitalmente il sistema appone un riferimento temporale utilizzato per il consolidamento in accordo con il DPCM 13/01/2004, art. 39 comma 1.

Il consolidamento del valore probatorio si completa con l'archiviazione (e successiva conservazione) delle copie delle CRL di pertinenza pubblicate prima e dopo l'istante identificato dal *timestamp* associato al documento, in modo da poter dimostrare con la massima attendibilità che il documento era valido in quel preciso momento. Il sistema prevede infatti una specifica sequenza di volumi di conservazione destinata a contenere (e conservare a lungo termine) le copie delle CRL di pertinenza di tutti i documenti elaborati, così da risultare indipendente – in sede di prova – dalle copie delle CRL archiviate dai certificatori.

I documenti per cui sia completato il ciclo di verifica e consolidamento vengono automaticamente inoltrati al sottosistema di conservazione.

2.3 Gestione firma digitale multipla

Il sistema è in grado di gestire le firme multiple in piena ottemperanza a quanto stabilito dal Codice dell'amministrazione digitale (decreto legislativo 5 marzo 2005 n. 82) e nella deliberazione CNIPA n. 34/2006 "Regole tecniche per la definizione del profilo di busta crittografica per la firma digitale in linguaggio XML".

La realizzazione delle firme multiple congiunte (altrimenti dette contestuali o parallele) è caratterizzata dal fatto che esse sono apposte in modo indipendente sugli stessi dati di partenza. Nel caso particolare di XML Signature possiamo evidenziare che:

- all'elemento Signature corrisponde una sola firma e quindi un solo firmatario;
- gli elementi Reference che individuano il documento, ovvero i documenti, firmati, sono contenuti nell'elemento Signature e contengono anche il DigestValue di ciascuno di essi.

L'attributo Signature ID deve essere valorizzato e deve essere univoco per ciascun firmatario.

2.4 Gestione marca temporale

Nella selezione della TSA, è importante selezionare un certificatore accreditato al comitato tecnico di interoperabilità di Assocertificatori.

Tale appartenenza garantisce che i documenti firmati e/o marcati rilasciati da questa TSA godano della interoperabilità tra i diver-

si certificatori accreditati.

Nello specifico l'interoperabilità garantisce che un oggetto firmato e/o marcato emesso dal certificatore venga correttamente verificato da un qualsiasi altro certificatore accreditato ed iscritto ad Assocertificatori e viceversa.

Il processo di verifica del sistema deve essere sempre in grado di:

- verificare l'integrità del documento firmato;
- dare la possibilità di estrarre il documento originale firmato;
- verificare la credibilità della Certification Authority che ha emesso il certificato del firmatario il documento;
- verificare la validità temporale del certificato del firmatario;
- verificare lo stato di validità del certificato del firmatario (accesso alla lista CRL/CSL del certificatore che ha emesso il certificato del firmatario il documento).

Al fine di assicurare l'interoperabilità, il CNIPA nel corso del 2005 ha emanato la deliberazione n. 4/2005 del 17 febbraio 2005, recante le "Regole per il riconoscimento e la verifica del documento informatico".

Ulteriormente all'interno del decreto del Presidente del Consiglio dei Ministri 13 gennaio 2004 "Regole tecniche per la formazione, la trasmissione, la conservazione, la duplicazione, la riproduzione e la validazione, anche temporale, dei documenti informatici".

In rispondenza nello specifico all'art. 45 sono riportate le informazioni che sono contenute nella marca temporale:

- identificativo dell'emittente;
- numero di serie della marca temporale;
- algoritmo di sottoscrizione della marca temporale;

- identificativo del certificato relativo alla chiave di verifica della marca;
- data ed ora di generazione della marca;
- identificazione dell'algoritmo di hash utilizzato per generare l'impronta dell'evidenza informatica sottoposta a validazione temporale;
- valore dell'impronta dell'evidenza informatica.

Le marche temporali vengono richieste tramite adapter che, a fronte della generazione dell'impronta, producono una richiesta verso il sistema di marcatura temporale.

L'operazione viene eseguita secondo la seguente procedura:

- l'impronta del documento viene inviata al servizio di marcatura temporale; l'impronta costituisce un riferimento certo al testo originale ma non ne consente la ricostruzione, in questo modo la marcatura può essere effettuata senza compromettere la "confidenzialità" del documento;
- il servizio di marcatura aggiunge all'impronta ricevuta la data e l'ora, ottenendo una "impronta marcata";
- l'impronta marcata è registrata con la chiave segreta del servizio, ottenendo la marca temporale da cui è possibile recuperare, mediante la chiave pubblica del servizio, l'impronta del documento, la data e l'ora della sua generazione;
- la marca temporale viene inviata al richiedente il quale la allega al documento.

2.5 Gestione dispositivi hardware di firma massiva (HSM)

Il sistema deve essere in grado di gestire i dispositivi di firma massiva utilizzati dalle CA conformi alla normativa vigente (ad es. i dispositivi HSM del tipo AEP Sureware Keyper Professional).

Su questi dispositivi vengono create le coppie di chiavi di sottoscrizione dei titolari che apporranno in modalità automatica e massiva la firma digitale ai documenti.

Per garantire il requisito della ridondanza in favore di un'alta affidabilità sono gestiti due dispositivi (con le medesime caratteristiche): per ciascun dispositivo ogni titolare avrà una diversa coppia di chiavi, così come richiesto dalla normativa vigente in tema di firma digitale.

Le principali caratteristiche sono:

- possibilità di gestione di chiavi RSA fino a 4.096 bit di lunghezza;
- algoritmi di hash SHA1- SHA 2 e MD-5 implementati in hardware;
- performance oltre 1.200 transazione al secondo (con chiavi di lunghezza pari a 1.024 bit);
- gestione di amministrazione del dispositivo con presenza contemporanea di due operatori autenticati tramite smart card;
- dispositivo di rete per l'utilizzo dello stesso da diversi sistemi;
- gestione di 32 accessi concorrenti.

2.6 Gestione processo di riversamento

Il processo di riversamento diretto consente di trasferire uno o più documenti conservati da un supporto ottico di memorizzazione ad un altro, senza alterarne la rappresentazione digitale (utilizzato ad esempio per creare copie di backup dei supporti di memorizzazione).

Il riversamento diretto può essere realizzato liberamente dal responsabile della conservazione, in quanto la delibera non prevede, al riguardo, specifiche prescrizioni formali.

Esso consiste nel trasferimento di uno o più documenti conservati da un supporto di memorizzazione a un altro, senza modificare la loro rappresentazione informatica. Si tratta, ad esempio, della generazione di copie di sicurezza da parte del responsabile della conservazione, come disposto dall'articolo 5, comma 1, della delibera.

Con il riversamento diretto viene effettuata quella che in termini tecnici viene definita clonazione del supporto, ossia, viene generato un supporto identico sia nel contenuto che nella rappresentazione dei file.

Il processo di riversamento sostitutivo consente di trasferire uno o più documenti conservati da un supporto ottico di memorizzazione ad un altro, modificandone la rappresentazione digitale.

Questo processo è necessario per garantire l'aggiornamento tecnologico dell'archivio quando non sia possibile o conveniente mantenere nel tempo il formato della rappresentazione digitale dei documenti.

L'operazione è necessaria per ottemperare l'obbligo di esibizione dei documenti conservati nel sistema.

A differenza del riversamento diretto, il riversamento sostitutivo è espressamente disciplinato dalla delibera, con l'articolo 3, comma 2, per quanto riguarda i documenti informatici, e con l'articolo 4, comma 4, per quanto concerne i documenti analogici.

Il riversamento sostitutivo, nel trasferimento di uno o più documenti conservati da un supporto di memorizzazione a un altro, modifica la rappresentazione informatica del suo contenuto. Si tratta, in questo caso, di un aggiornamento tecnologico dell'archivio informatico, in quanto non è più conveniente mantenere nel tempo il formato di rappresentazione digitale dei documenti

originariamente conservati.

Il processo si conclude con l'apposizione, sull'insieme dei documenti o su una evidenza informatica contenente una o più impronte dei documenti, della firma digitale da parte del responsabile della conservazione e della marca temporale.

Qualora il riversamento riguardi documenti analogici originali unici (documenti dei quali non è possibile risalire al contenuto degli stessi attraverso altre scritture di cui sia obbligatoria la conservazione), il processo si conclude con l'ulteriore apposizione della firma digitale da parte di un pubblico ufficiale e della marca temporale.

Si precisa che dal 1° gennaio 2006 le disposizioni in materia di documento informatico sottoscritto ai sensi dell'articolo 10 del Dpr 445/2000 sono state abrogate dal decreto legislativo 82/2005 "Codice dell'amministrazione digitale", di conseguenza la definizione di documento informatico si ricava dalle disposizioni recate dall'articolo 20 del sopra citato codice. Pertanto, essendo i documenti informatici assimilati a quelli digitali generati da documenti analogici originali unici, il processo di riversamento sostitutivo di tali documenti si perfeziona con l'apposizione del riferimento temporale e della firma digitale da parte di un pubblico ufficiale.

Si evidenzia, infine, che il documento informatico sottoscritto con firma digitale ha l'efficacia prevista dall'articolo 2702 cc, ovvero fa piena prova, fino a querela di falso.

2.7 Gestione processo di esibizione

Per esibizione del documento si intende la visualizzazione di un documento o di un intero supporto con la quale si fornì-

scono tutti i dettagli che permettono di verificare il processo di conservazione.

È possibile verificare immediatamente l'integrità e la validità del documento e del supporto che lo contiene e contestualmente possono essere evidenziati gli estremi della certificazione inclusa la firma del responsabile e le informazioni sulla marca temporale apposta.

Dall'interfaccia di esibizione è possibile, inoltre, richiedere la materializzazione su CD e DVD dell'intero supporto con tutte le informazioni che ne permettono le verifiche formali.

La console del sistema è a disposizione del responsabile della conservazione e del personale addetto alle ispezioni.

2.8 Virgilio, la soluzione per la Conservazione Sostitutiva a lungo termine

Virgilio è un innovativo sistema di Conservazione Sostitutiva creato da Siav per gestire in modo semplice e sicuro la Conservazione Sostitutiva a lungo termine dei documenti informatici, passando dal concetto di supporto fisico sostitutivo a quello più evoluto ed efficiente di archivio unico sostitutivo.

Virgilio integra una serie di servizi specifici di monitoraggio dello stato fisico e logico dell'archivio, effettuando, per ogni documento conservato, una continua verifica di caratteristiche come la leggibilità, l'integrità, il valore legale, l'obsolescenza del formato e la possibilità di applicare la procedura di scarto d'archivio.

Virgilio ottempera a tutti i prerequisiti e le funzionalità specificate nelle pagine precedenti.

3. Definizione di uno standard di interoperabilità per la gestione dei volumi di conservazione

L'opportunità offerta dai documenti digitali in termini di efficienza operativa dei processi di business correlati e di semplicità di fruizione, condivisione e distribuzione dell'informazione stanno spingendo sempre più le aziende private e gli enti pubblici a dotarsi di strumenti tecnologici in grado di supportare il cambiamento verso l'era digitale.

Ai classici archivi cartacei si affiancano oggi, ben più ricchi di contenuti e largamente condivisi, gli archivi digitali dove confluiscono le forme documentali proprie dell'Information Technology quali ad esempio i messaggi di posta elettronica, i documenti Office, i documenti PDF, le pagine html, i filmati oltre ovviamente ai documenti cartacei digitalizzati.

Il Codice dell'amministrazione digitale rappresenta una pietra miliare in questo percorso evolutivo in quanto definisce in modo chiaro e preciso come i nuovi strumenti tecnologici (la firma digitale, la posta elettronica certificata e la marca temporale) possono essere utilizzati per dare valore legale ai documenti elettronici (documento informatico) definendo al contempo le regole per una corretta conservazione degli stessi.

La possibilità di conservare in forma sostitutiva i documenti analogici (cartacei, microfilm) trasformandoli in documenti digitali a valore legale consente inoltre di gestire in modo ottimale gli archivi misti, composti cioè da documenti che in origine potevano essere sia cartacei che elettronici.

Su questo punto tuttavia, anche se la norma appare chiara, ci sono ancora troppi dubbi e incertezze legate ad esempio al valore giuridico di una firma autografa su un documento digitalizzato: il ripudio della firma autografa richiede in sede di giudizio

una perizia calligrafica che risulta essere meno attendibile su un documento digitalizzato rispetto a un documento cartaceo (a causa della perdita delle informazioni legate alla profondità del segno lasciato dalla penna sul foglio).

Queste incertezze sono di fatto un freno al processo di conservazione sostitutiva dei documenti cartacei, che ad oggi viene preso in considerazione solo per documenti che abbiano un “basso” valore legale.

Le soluzioni che si limitano a gestire la sola fase di produzione dei supporti fisici a normativa CNIPA appaiono oggi inadeguate nel momento in cui si voglia affrontare un modello multiaziendale complesso dove siano presenti tipologie documentali con flussi di lavorazione differenti (documenti sanitari, fiscali, lavoro).

L'esigenza è quella di passare dalla gestione di un archivio di supporti fisici a quella di un repository unico virtuale, l'archivio sostitutivo, in grado di conservare formati documentali differenti, garantendo sia un elevato livello di affidabilità e controllo che una semplificazione di tutte le attività amministrative e di maintenance.

Concentrandosi sugli aspetti tecnici e procedurali, la normativa sulla conservazione sostitutiva, fa oggi riferimento alle regole tecniche emanate dal CNIPA con la deliberazione 11/2004, che definisce le procedure da rispettare per la realizzazione di supporti sostitutivi a norma di legge.

Nell'allegato tecnico, al paragrafo 4, si precisa che il supporto di conservazione non deve essere necessariamente ottico (CD/DVD) in quanto il processo di conservazione consente di produrre un “sigillo informatico”, detto file di chiusura, a garanzia dell'immodificabilità dei documenti conservati.

La norma non definisce una specifica tecnica di standardizzazio-

ne per la produzione del file di chiusura, in quanto si limita ad indicare che il processo di certificazione “termina con l'apposizione, sull'insieme dei documenti o su una evidenza informatica contenente una o più impronte dei documenti o di insiemi di essi, del riferimento temporale e della firma digitale da parte del responsabile della conservazione”.

Questo porta come immediata conseguenza che produttori di soluzioni informatiche per la gestione del procedimento di conservazione sostitutiva utilizzano schemi differenti per la creazione del file di chiusura e del volume di conservazione (supporto sostitutivo).

Tutte le informazioni necessarie per verificare l'integrità del volume di conservazione ovviamente devono essere presenti all'interno del file di chiusura, ma di fatto si rendono molto difficili le operazioni di verifica a meno di non utilizzare per tale attività la stessa soluzione applicativa che ha generato il supporto, che di norma prevede anche un modulo di verifica.

Tale approccio può portare al caso limite in cui si può riscontrare che applicazioni sviluppate in tempi successivi dalla stessa software house non garantiscono la retrocompatibilità nelle operazioni di verifica.

La mancanza di uno schema standard da utilizzare per la generazione del file di chiusura è quindi un problema concreto che deve essere gestito dal responsabile della conservazione, soprattutto nel caso debba garantire il valore legale dei documenti conservati per periodi di tempo molto lunghi.

Lo stesso problema, relativo però ai formati documentali, è già stato affrontato dalla comunità internazionale definendo lo standard ISO-19005-1 “Document Management Electronic document file format for long-term preservation” che ha portato

alla definizione del formato aperto PDF/A, riconosciuto oramai come la scelta più adatta per la conservazione di documenti informatici.

La definizione di un formato standard per il file di chiusura e, più in generale, dei supporti di conservazione è sicuramente un compito che spetta al CNIPA, anche se non è da escludere che il mercato possa produrre delle proposte al fine di definire uno standard *de facto*.

Siav, all'interno della propria piattaforma di Conservazione Sostitutiva Enterprise denominata Virgilio, ha definito una struttura per il file di chiusura basata sullo standard XML che consente una chiara e semplice interpretazione delle informazioni relative ad ogni singolo documento conservato all'interno del supporto.

La proposta è quella di definire all'interno del file di chiusura un blocco XML di testata specifico per il volume di conservazione e per ogni singolo documento un blocco XML descrittivo composto da quattro distinte sezioni:

- Sezione 1: posizione fisica (relativa/assoluta)
- Sezione 2: metadati di classificazione (tipizzati)
- Sezione 3: informazioni aggiuntive (formato, firme digitali, marcatura temporale)
- Sezione 4: impronta

Oltre alla struttura base del blocco è necessario definire un elenco di tag utilizzabili, descrivendo in modo chiaro, per ciascuno di essi, la tipologia di informazioni che verrà gestita al loro interno, al fine di consentire una chiara ed univoca interpretazione delle stesse.

Un esempio di struttura XML relativa ad un singolo documento potrebbe essere il seguente:

```
<documento>
<Sezione1>
<DocPath RelPath="5468BD08-9B2C-441C-89F3-FB33BA216879"/>
</Sezione1>
<Sezione2>
<chiave nome="Rag.Sociale" tipo="xs:string">ROSSI GIORGIO </chiave>
<chiave nome="Cod.Fiscale" tipo="xs:string" />
<datadocumento>2004-10-31T00:00:00Z</datadocumento>
</Sezione2>
<Sezione3>
<formatonativo> PDF/A </formatonativo>
</Sezione3>
<Sezione4>
<impronta>37B8B54BB4762BEA9CFB93C880097959106C25AE</impronta>
</Sezione4>
</documento>
```

L'aderenza delle applicazioni di conservazione sostitutiva ad uno "standard" come quello illustrato potrebbe consentire ai diversi prodotti presenti sul mercato di interoperare, rendendo di fatto indipendenti la fase di creazione da quella di verifica e maintenance.

La garanzia che i supporti sostitutivi prodotti possano essere verificati e mantenuti validi a prescindere dall'applicazione che li ha generati è sicuramente uno degli aspetti più importanti per il responsabile della conservazione il quale, in base alla delibera CNIPA 11/04, "mantiene e rende accessibile un archivio del software dei programmi in gestione nelle eventuali diverse versioni;" (Art. 5-c) ed inoltre "verifica periodicamente, con cadenza

non superiore a cinque anni, l'effettiva leggibilità dei documenti conservati provvedendo, se necessario, al riversamento diretto o sostitutivo del contenuto dei supporti" (Art. 5-h).

A tale scopo, tra le attività di ricerca avviate dal Comitato Scientifico Siav una riguarda proprio la standardizzazione del formato per il file di chiusura, che porterà in tempi brevi alla redazione della documentazione tecnica necessaria a definire una proposta concreta ed aperta a tutti gli operatori di mercato.

Siav auspica che tale proposta possa essere la base di partenza sulla quale organizzare un tavolo di lavoro tra i principali operatori e le associazioni del settore, avente come obiettivo la definizione di una proposta concreta da presentare al CNIPA per la definizione di uno standard ufficiale.

I vantaggi di tale adozione sono evidenti: trasparenza e semplicità nell'accesso ai dati indipendentemente dall'azienda che ha realizzato il prodotto e le procedure di conservazione sostitutiva, oltre alla possibilità di aziende ed enti di interoperare e quindi di non essere dipendenti dal fornitore delle procedure, dando maggiore valore allo spirito della norma.

Appendice 1

Supporti di memorizzazione a norma per la conservazione

Come è noto, la normativa ufficiale sulle modalità per memorizzare i documenti informatici fa riferimento ai supporti ottici limitando pertanto, secondo una rigorosa interpretazione letterale, a questa tecnologia e soltanto a questa la possibilità di essere utilizzata per registrare le informazioni in forma "permanente".

L'origine di questa disposizione sta nella Legge 24 dicembre 1993, n. 537, *Interventi correttivi di finanza pubblica*, che all'articolo 2, comma 15, dispone: "Gli obblighi di conservazione e di esibizione di documenti, per finalità amministrative e probatorie, previsti dalla legislazione vigente, si intendono soddisfatti anche se realizzati mediante supporto ottico³⁴ purché le procedure utilizzate siano conformi a regole tecniche dettate [...] dall'Autorità per l'informatica nella pubblica amministrazione".

Giova ricordare che nei primi anni '90 sembrava che la tecnologia ottica fosse destinata a dominare, in futuro, il mercato delle archiviazioni di massa mentre le tecnologie magnetiche tradizionali si avviavano verso il tramonto. Come è ben noto, a partire dalla metà degli anni '90 la realtà prese strade completamente diverse.

Traccia di questo cambiamento di prospettiva si ha nel *Testo Unico delle disposizioni legislative e regolamentari in materia di documentazione amministrativa* (DPR 28 dicembre 2000, n.

³⁴ Il supporto ottico introduce qui, per la prima volta, la "forma digitale" di un documento, in aggiunta alla tradizionale forma cartacea alla quale si era già affiancata quella su microfilm.

445), che, riprendendo nell'articolo 6 il concetto espresso nella norma del 1993 sopra riportata, stabilisce:

“1. Le pubbliche amministrazioni ed i privati hanno facoltà di sostituire, a tutti gli effetti, i documenti dei propri archivi, le scritture contabili, la corrispondenza e gli altri atti di cui per legge o regolamento è prescritta la conservazione, con la loro riproduzione su supporto fotografico, su supporto ottico o con altro mezzo idoneo a garantire la conformità dei documenti agli originali.

2. Gli obblighi di conservazione ed esibizione dei documenti di cui al comma 1 si intendono soddisfatti, sia ai fini amministrativi che probatori, anche se realizzati su supporto ottico (*si noti che manca l'alternativa “o con altro mezzo”, N.d.R*) quando le procedure utilizzate sono conformi alle regole tecniche dettate dall'Autorità per l'informatica nella pubblica amministrazione”.

Incidentalmente, si deve notare che il comma 1 contiene un grossolano errore concettuale: infatti non è certamente il supporto fisico che garantisce la conformità dei documenti agli originali. Il supporto può solo assicurare il mantenimento dell'integrità della registrazione per una certa durata di tempo, il che è cosa ben diversa dalla conformità all'originale.

A seguito di questa norma l'AIPA ha dettato il 13 dicembre 2001 la deliberazione n. 42 che all'articolo 8 stabilisce: “Tenuto conto dell'evoluzione tecnologica e della disciplina dettata dal DPR 28 dicembre 2000, n. 445, è data facoltà alle pubbliche amministrazioni e ai privati, ove non ostino altre motivazioni, di utilizzare, nel processo di conservazione digitale, un qualsiasi supporto di memorizzazione, anche non ottico, comunque idoneo a garantire la conformità dei documenti agli originali, nel rispetto delle modalità previste dalla presente deliberazione”.

Si noti che anche in questo testo è rimasto l'errore sopra segnalato di “conformità agli originali” come proprietà attribuita al supporto.

Il concetto è stato successivamente ripreso in modo identico dalla deliberazione CNIPA n. 11 del 19 febbraio 2004, che aggiornava, sostituendola, la deliberazione n. 42 sopra citata.

L'autorizzazione all'utilizzo di un qualsiasi tipo di supporto di memorizzazione viene ribadito nelle *Note esplicative* che accompagnano la deliberazione n. 11, che, al paragrafo 4, così si esprimono:

“La deliberazione autorizza l'utilizzazione di un qualsiasi tipo di supporto di memorizzazione che consenta la registrazione mediante la tecnologia laser; quindi, non soltanto dischi ottici WORM e CD-R, ma anche magneto-ottici e DVD. Inoltre, tenuto conto di quanto previsto nel già citato testo unico circa la possibilità d'impiego ai fini della conservazione di altro mezzo idoneo a garantire la conformità dei documenti agli originali e in considerazione dell'evoluzione tecnologica nel frattempo avvenuta, la deliberazione consente (art. 8) di utilizzare un qualsiasi altro supporto di memorizzazione, oltre quelli a tecnologia laser, se non ostino particolari motivazioni e comunque nel rispetto delle regole tecniche previste dalla deliberazione stessa. Si ritiene, infatti, che i tecnicismi della firma digitale e del riferimento temporale siano idonei a garantire l'integrità del documento nel processo di sua conservazione, qualunque sia il tipo di supporto di memorizzazione, anche se diverso da quello ottico. Soltanto specifiche motivazioni potrebbero suggerire o richiedere l'impiego di una particolare tipologia di supporto”.

È stato però contestato il potere dell'AIPA, e di conseguenza anche del CNIPA, di disporre varianti sulla tecnologia, in quanto l'articolo 6 del citato Testo Unico al comma 3 specifica: “I limiti

e le modalità tecniche della riproduzione e dell'autenticazione dei documenti di cui al comma 1, su supporto fotografico o con altro mezzo tecnico idoneo a garantire la conformità agli originali, sono stabiliti con decreto del Presidente del Consiglio dei Ministri”.

Pertanto, sarebbe stato necessario un Decreto del Presidente del Consiglio dei Ministri, e non una “semplice” deliberazione del CNIPA, per disporre diversamente sulla tecnologia da adottare. In altre parole, il CNIPA può dettare regole tecniche sull'archiviazione ottica, ma non sull'archiviazione basata sulla fotografia o su altra tecnologia.

Il Codice dell'amministrazione digitale (Decreto legislativo 7 marzo 2005, n. 82, e seg.) ha poi confermato, ma solo in parte, l'articolo 6 del Testo Unico: infatti nell'articolo 43, comma 1, si legge: “I documenti degli archivi, le scritture contabili, la corrispondenza ed ogni atto, dato o documento di cui è prescritta la conservazione per legge o regolamento, ove riprodotti su supporti informatici sono validi e rilevanti a tutti gli effetti di legge, se la riproduzione sia effettuata in modo da garantire la conformità dei documenti agli originali e la loro conservazione nel tempo, nel rispetto delle regole tecniche stabilite ai sensi dell'articolo 71”.

In questo testo non si fa più riferimento a nessuna tecnologia in particolare, ma si rimanda alle nuove regole tecniche da emanare.

Però, finché queste non verranno pubblicate, restano in vigore le attuali, che, nel caso specifico, sono contenute nella deliberazione 11 del CNIPA, con le perplessità sopra descritte.

Il problema non è banale: poiché di fatto l'unica tecnologia oggi sicuramente ammessa per conservare i documenti è quella ottica, se un'Amministrazione dovesse incorrere, per un motivo

qualunque, nel caso disgraziato di una perdita totale di dati conservati con tecnologie diverse da quelle ottiche, il responsabile della struttura informatica potrebbe essere chiamato a rispondere del danno per non avere seguito le norme. A poco potrebbe valere, davanti al magistrato, il ricorso alla deliberazione CNIPA, se ritenuta affetta da eccesso di delega in termini di poteri regolatori.

Resta il fatto, ad ammorbidire l'affermazione precedente, che il dettato dell'articolo 8 di cui sopra è stato adottato da molti dei maggiori fornitori di servizi di Conservazione Sostitutiva, confortati dalla realtà tecnologica che, come illustrato in questo testo, poggia su un corretto ed appropriato uso delle memorie magnetiche.

Appendice 2

Deliberazione CNIPA n. 11 del 19 febbraio 2004

(Gazzetta Ufficiale - Serie Generale n. 57 del 9-3-2004)

Regole tecniche per la riproduzione e conservazione di documenti su supporto ottico idoneo a garantire la conformità dei documenti agli originali - Art. 6, commi 1 e 2, del testo unico delle disposizioni legislative e regolamentari in materia di documentazione amministrativa, di cui al decreto del Presidente della Repubblica 28 dicembre 2000, n. 445.

Il Collegio

Visto l'art. 4, comma 1, del decreto legislativo 12 febbraio 1993, n. 39, così come sostituito dall'art. 176, comma 3, del decreto legislativo 2003, n. 196, che istituisce il Centro nazionale per l'informatica nella pubblica amministrazione;

Visto l'art. 6, comma 1, del decreto del Presidente della Repubblica 28 dicembre 2000, n. 445, il quale prevede che le pubbliche amministrazioni ed i privati hanno facoltà di sostituire, a tutti gli effetti, i documenti dei propri archivi, le scritture contabili, la corrispondenza e gli altri atti di cui per legge o regolamento è prescritta la conservazione, con la loro riproduzione su supporto fotografico, su supporto ottico o con altro mezzo idoneo a garantire la conformità dei documenti agli originali;

Visto l'art. 6, comma 2, del citato decreto del Presidente della Repubblica 28 dicembre 2000, n. 445, il quale prevede che gli obblighi di conservazione ed esibizione dei documenti di cui al comma 1 si intendono soddisfatti, ai fini sia amministrativi sia probatori, anche se realizzati su supporto ottico quando le

procedure utilizzate sono conformi alle regole tecniche dettate dall'Autorità' per l'informatica nella pubblica amministrazione;

Vista la deliberazione AIPA n. 42 del 13 dicembre 2001, con la quale sono state dettate le regole tecniche per la riproduzione e conservazione di documenti su supporto ottico idoneo a garantire la conformità dei documenti agli originali;

Ritenuto di dover procedere alla revisione prevista dall'art. 10 della citata deliberazione n. 42/2001 al fine dell'adeguamento alle esigenze di rinnovamento tecnologico e che, pertanto, è necessario provvedere all'adozione di una nuova deliberazione che sostituisca integralmente la detta delibera n. 42/2001;

Delibera:

La presente deliberazione, che sostituisce la deliberazione n. 42 del 13 dicembre 2001, trova applicazione dal giorno successivo alla sua pubblicazione nella Gazzetta Ufficiale della Repubblica italiana.

Articolo 1

Definizioni

1. Ai fini della presente deliberazione si intende per:

- a) documento: rappresentazione informatica o in formato analogico di atti, fatti e dati intelligibili direttamente o attraverso un processo di elaborazione elettronica;
- b) documento analogico: documento formato utilizzando una grandezza fisica che assume valori continui, come le tracce su carta (esempio: documenti cartacei), come le immagini su film (esempio: pellicole mediche, microfiche,

microfilm), come le magnetizzazioni su nastro (esempio: cassette e nastri magnetici audio e video). Si distingue in documento originale e copia;

- c) documento analogico originale: documento analogico che può essere unico oppure non unico se, in questo secondo caso, sia possibile risalire al suo contenuto attraverso altre scritture o documenti di cui sia obbligatoria la conservazione, anche se in possesso di terzi;
- d) documento informatico: la rappresentazione informatica di atti, fatti o dati giuridicamente rilevanti;
- e) supporto ottico di memorizzazione: mezzo fisico che consente la memorizzazione di documenti informatici mediante l'impiego della tecnologia laser (quali, ad esempio, dischi ottici, magneto-ottici, DVD);
- f) memorizzazione: processo di trasposizione su un qualsiasi idoneo supporto, attraverso un processo di elaborazione, di documenti analogici o informatici, anche sottoscritti ai sensi dell'art. 10, commi 2 e 3, del decreto del Presidente della Repubblica 28 dicembre 2000, n. 445, così come modificato dall'art. 6 del decreto legislativo 23 gennaio 2002, n. 10;
- g) archiviazione elettronica: processo di memorizzazione, su un qualsiasi idoneo supporto, di documenti informatici, anche sottoscritti, così come individuati nella precedente lettera f), univocamente identificati mediante un codice

- di riferimento, antecedente all'eventuale processo di conservazione;
- h) documento archiviato: documento informatico, anche sottoscritto, così come individuato nella precedente lettera f), sottoposto al processo di archiviazione elettronica;
 - i) conservazione sostitutiva: processo effettuato con le modalità di cui agli articoli 3 e 4 della presente deliberazione;
 - l) documento conservato: documento sottoposto al processo di conservazione sostitutiva;
 - m) esibizione: operazione che consente di visualizzare un documento conservato e di ottenerne copia;
 - n) riversamento diretto: processo che trasferisce uno o più documenti conservati da un supporto ottico di memorizzazione ad un altro, non alterando la loro rappresentazione informatica. Per tale processo non sono previste particolari modalità;
 - o) riversamento sostitutivo: processo che trasferisce uno o più documenti conservati da un supporto ottico di memorizzazione ad un altro, modificando la loro rappresentazione informatica. Per tale processo sono previste le modalità descritte nell'art. 3, comma 2, e nell'art. 4, comma 4, della presente deliberazione;
 - p) riferimento temporale: informazione, contenente la data e l'ora, che viene associata ad uno o più documenti informatici;

- q) pubblico ufficiale: il notaio, salvo quanto previsto dall'art. 5, comma 4 della presente deliberazione e nei casi per i quali possono essere chiamate in causa le altre figure previste dall'art. 18, comma 2, del decreto del Presidente della Repubblica 28 dicembre 2000, n. 445;
- r) evidenza informatica: una sequenza di simboli binari (bit) che può essere elaborata da una procedura informatica;
- s) impronta: la sequenza di simboli binari (bit) di lunghezza predefinita generata mediante l'applicazione alla prima di una opportuna funzione di hash;
- t) funzione di hash: una funzione matematica che genera, a partire da una generica sequenza di simboli binari (bit), una impronta in modo tale che risulti di fatto impossibile, a partire da questa, determinare una sequenza di simboli binari (bit) che la generi, ed altresì risulti di fatto impossibile determinare una coppia di sequenze di simboli binari per le quali la funzione generi impronte uguali;
- u) firma digitale: così come definita all'art. 1, comma 1, lettera n), del decreto del Presidente della Repubblica 28 dicembre 2000, n. 445.

Articolo 2

Obblighi di conservazione sostitutiva

1. Gli obblighi di conservazione sostitutiva dei documenti, previsti dalla legislazione vigente sia per le pubbliche

amministrazioni sia per i privati, sono soddisfatti a tutti gli effetti, fatto salvo quanto indicato dall'art. 7, qualora il processo di conservazione venga effettuato con le modalità di cui agli articoli 3 e 4.

2. I documenti informatici, anche sottoscritti, così come individuati nell'art. 1, lettera f), possono essere archiviati elettronicamente prima di essere sottoposti al processo di conservazione. Per l'archiviazione elettronica non sussistono gli obblighi di cui alla presente deliberazione.

Articolo 3

Conservazione sostitutiva di documenti informatici

1. Il processo di conservazione sostitutiva di documenti informatici, anche sottoscritti, così come individuati nell'art. 1, lettera f), e, eventualmente, anche delle loro impronte, avviene mediante memorizzazione su supporti ottici e termina con l'apposizione, sull'insieme dei documenti o su una evidenza informatica contenente una o più impronte dei documenti o di insiemi di essi, del riferimento temporale e della firma digitale da parte del responsabile della conservazione che attesta il corretto svolgimento del processo.
2. Il processo di riversamento sostitutivo di documenti informatici conservati avviene mediante memorizzazione su altro supporto ottico e termina con l'apposizione sull'insieme dei documenti o su una evidenza informatica contenente una o più impronte dei documenti o di insiemi di essi del riferimento temporale e della firma digitale da parte del responsabile della conservazione che attesta il corretto svolgimento del processo.

Qualora il processo riguardi documenti informatici sottoscritti, così come individuati nell'art. 1, lettera f), è inoltre richiesta l'apposizione del riferimento temporale e della firma digitale, da parte di un pubblico ufficiale, per attestare la conformità di quanto riversato al documento d'origine.

Articolo 4

Conservazione sostitutiva di documenti analogici

1. Il processo di conservazione sostitutiva di documenti analogici avviene mediante memorizzazione della relativa immagine direttamente sui supporti ottici, eventualmente, anche della relativa impronta, e termina con l'apposizione, sull'insieme dei documenti o su una evidenza informatica contenente una o più impronte dei documenti o di insiemi di essi, del riferimento temporale e della firma digitale da parte del responsabile della conservazione che attesta così il corretto svolgimento del processo.
2. Il processo di conservazione sostitutiva di documenti analogici originali unici si conclude con l'ulteriore apposizione del riferimento temporale e della firma digitale da parte di un pubblico ufficiale per attestare la conformità di quanto memorizzato al documento d'origine.
3. La distruzione di documenti analogici, di cui è obbligatoria la conservazione, è consentita soltanto dopo il completamento della procedura di conservazione sostitutiva, fatto salvo quanto previsto al comma 4 dell'art. 6 del decreto del Presidente della Repubblica 28 dicembre 2000, n. 445.

4. Il processo di riversamento sostitutivo di documenti analogici conservati avviene mediante memorizzazione su altro supporto ottico. Il responsabile della conservazione, al termine del riversamento, ne attesta il corretto svolgimento con l'apposizione del riferimento temporale e della firma digitale sull'insieme dei documenti o su una evidenza informatica contenente una o più impronte dei documenti o di insiemi di essi. Qualora il processo riguardi documenti originali unici di cui al comma 2, è richiesta l'ulteriore apposizione del riferimento temporale e della firma digitale da parte di un pubblico ufficiale per attestare la conformità di quanto riversato al documento d'origine.

Articolo 5

Responsabile della conservazione

1. Il responsabile del procedimento di conservazione sostitutiva:
 - a) definisce le caratteristiche e i requisiti del sistema di conservazione in funzione della tipologia dei documenti (analogici o informatici) da conservare, della quale tiene evidenza. Organizza conseguentemente il contenuto dei supporti ottici e gestisce le procedure di sicurezza e di tracciabilità che ne garantiscono la corretta conservazione, anche per consentire l'esibizione di ciascun documento conservato;
 - b) archivia e rende disponibili, con l'impiego di procedure elaborative, relativamente ad ogni supporto di memorizzazione utilizzato, le seguenti informazioni:

- 1) descrizione del contenuto dell'insieme dei documenti;
- 2) estremi identificativi del responsabile della conservazione;
- 3) estremi identificativi delle persone eventualmente delegate dal responsabile della conservazione, con l'indicazione dei compiti alle stesse assegnati;
- 4) indicazione delle copie di sicurezza;
- c) mantiene e rende accessibile un archivio del software dei programmi in gestione nelle eventuali diverse versioni;
- d) verifica la corretta funzionalità del sistema e dei programmi in gestione;
- e) adotta le misure necessarie per la sicurezza fisica e logica del sistema preposto al processo di conservazione sostitutiva e delle copie di sicurezza dei supporti di memorizzazione;
- f) richiede la presenza di un pubblico ufficiale nei casi in cui sia previsto il suo intervento, assicurando allo stesso l'assistenza e le risorse necessarie per l'espletamento delle attività al medesimo attribuite;
- g) definisce e documenta le procedure di sicurezza da rispettare per l'apposizione del riferimento temporale;
- h) verifica periodicamente, con cadenza non superiore a cinque anni, l'effettiva leggibilità dei documenti conservati provvedendo, se necessario, al riversamento diretto o sostitutivo del contenuto dei supporti.

2. Il responsabile del procedimento di conservazione sostitutiva può delegare, in tutto o in parte, lo svolgimento delle proprie attività ad una o più persone che, per competenza ed esperienza, garantiscano la corretta esecuzione delle operazioni ad esse delegate.
3. Il procedimento di conservazione sostitutiva può essere affidato, in tutto o in parte, ad altri soggetti, pubblici o privati, i quali sono tenuti ad osservare quanto previsto dalla presente deliberazione.
4. Nelle amministrazioni pubbliche il ruolo di pubblico ufficiale è svolto dal dirigente dell'ufficio responsabile della conservazione dei documenti o da altri dallo stesso formalmente designati, fatta eccezione per quanto previsto dall'art. 3, comma 2, e dall'art. 4, commi 2 e 4, casi nei quali si richiede l'intervento di soggetto diverso della stessa amministrazione.

Articolo 6

Obbligo di esibizione

1. Il documento conservato deve essere reso leggibile in qualunque momento presso il sistema di conservazione sostitutiva e disponibile, a richiesta, su supporto cartaceo.
2. Il documento conservato può essere esibito anche per via telematica.
3. Qualora un documento conservato venga esibito su supporto cartaceo fuori dall'ambiente in cui è installato il sistema di conservazione sostitutiva, deve esserne dichiarata la conformità da parte di un pubblico ufficiale se si tratta di documenti per la cui conservazione è previsto il suo intervento.

Articolo 7

Procedure operative

1. A qualsiasi soggetto pubblico o privato che intenda avvalersi del processo di conservazione sostitutiva dei documenti è consentita l'adozione di accorgimenti e procedure integrative, nel rispetto di quanto stabilito nella presente deliberazione.
2. Le pubbliche amministrazioni comunicano preliminarmente al Centro nazionale per l'informatica nella pubblica amministrazione le procedure integrative che intendono adottare ai sensi del comma 1.

Articolo 8

Altri supporti di memorizzazione

1. Tenuto conto dell'evoluzione tecnologica e della disciplina dettata dal decreto del Presidente della Repubblica 28 dicembre 2000, n. 445, è data facoltà alle pubbliche amministrazioni e ai privati, ove non ostino particolari motivazioni, di utilizzare, nei processi di conservazione sostitutiva e di riversamento sostitutivo, un qualsiasi supporto di memorizzazione, anche non ottico, comunque idoneo a garantire la conformità dei documenti agli originali, nel rispetto delle modalità previste dalla presente deliberazione.

Articolo 9

Sistemi di conservazione preesistenti

1. Le regole tecniche dettate con le deliberazioni n. 15 del 28 luglio 1994, n. 24 del 30 luglio 1998 e n. 42 del

13 dicembre 2001 continuano ad applicarsi ai sistemi di conservazione sostitutiva già esistenti o in corso di acquisizione al momento della pubblicazione della presente deliberazione.

2. I documenti conservati in osservanza delle regole tecniche indicate al comma 1 possono essere riversati in un sistema di conservazione sostitutiva tenuto in conformità alle regole tecniche dettate con la presente deliberazione.

Appendice 3

Conservazione dei documenti originali unici

Di seguito sono riportati i commi 4 e 5 dell'articolo 23 del Codice dell'amministrazione digitale (Decreto legislativo 7 marzo 2005, n. 82), così come modificati³⁵ dal comma 12 dell'articolo 16 del decreto-legge 29 novembre 2008, n. 185, recante misure urgenti per il sostegno a famiglie, lavoro, occupazione e impresa e per ridisegnare in funzione anti-crisi il quadro strategico nazionale, convertito in legge dalla Legge 28 gennaio 2009, n. 2.

“4. Le copie su supporto informatico di qualsiasi tipologia di documenti analogici originali, formati in origine su supporto cartaceo o su altro supporto non informatico, sostituiscono ad ogni effetto di legge gli originali da cui sono tratte se la loro conformità all'originale è assicurata da chi lo detiene mediante l'utilizzo della propria firma digitale e nel rispetto delle regole tecniche di cui all'articolo 71.

5. Con decreto del Presidente del Consiglio dei Ministri possono essere individuate particolari tipologie di documenti analogici originali unici per

³⁵ Versione previgente:

Articolo 23 - Copie di atti e documenti informatici

... Omissis ...

4. Le copie su supporto informatico di documenti originali non unici formati in origine su supporto cartaceo o, comunque, non informatico sostituiscono, ad ogni effetto di legge, gli originali da cui sono tratte se la loro conformità all'originale è assicurata dal responsabile della conservazione mediante l'utilizzo della propria firma digitale e nel rispetto delle regole tecniche di cui all'articolo 71.

5. Le copie su supporto informatico di documenti, originali unici, formati in origine su supporto cartaceo o, comunque, non informatico sostituiscono, ad ogni effetto di legge, gli originali da cui sono tratte se la loro conformità all'originale è autenticata da un notaio o da altro pubblico ufficiale a ciò autorizzato, con dichiarazione allegata al documento informatico e asseverata secondo le regole tecniche stabilite ai sensi dell'articolo 71.

le quali, in ragione di esigenze di natura pubblicistica, permane l'obbligo della conservazione dell'originale analogico oppure, in caso di conservazione ottica sostitutiva, la loro conformità all'originale deve essere autenticata da un notaio o da altro pubblico ufficiale a ciò autorizzato con dichiarazione da questi firmata digitalmente ed allegata al documento informatico.”

Breve commento

Con questa prescrizione è stato fatto un importante passo avanti in tema d'innovazione.

Questa nuova formulazione consente di convertire in digitale (“in grado di sostituire ad ogni effetto di legge gli originali”) e conservare in tale formato (con la sola attestazione di conformità all'originale cartaceo, tramite l'apposizione della firma digitale, da parte “di chi lo detiene”³⁶) anche la maggior parte dei documenti originali unici.

Viene dunque meno l'esigenza dell'intervento del pubblico ufficiale; questa però resta per alcune tipologie di documenti, per i quali sono previste particolari esigenze di natura pubblicistica, che dovranno essere individuati da un apposito decreto del Presidente del Consiglio dei Ministri .

Si fa rilevare che il dispositivo è esecutivo e l'eventuale decreto sicuramente non indicherà tutti quei documenti, originali unici, risultato di accordi commerciali o contrattuali tra privati.

A questo punto, in attesa che vengano definite e pubblicate le regole tecniche sulla conservazione di documenti (analogici e informatici) di cui all'articolo 71 del Codice dell'amministrazione digitale, è da ritenersi superata, almeno per quei documenti per i quali non sono evidenti particolari esigenze di natura pubbli-

³⁶ Ben inteso si intende a pieno titolo e legalmente.

stica, la prescrizione di intervento del pubblico ufficiale prevista dalla deliberazione CNIPA 11/2004.

Per amore di obiettività è giusto, in conclusione, sottolineare che l'interpretazione qui fornita non è condivisa da altri, anche autorevoli interpreti delle norme in questione, che sostengono che l'articolo 23 del Codice dell'amministrazione digitale non si occupa di conservazione ma solo della formazione di copie di documenti.

Si fa, dunque, sempre più pressante la pubblicazione del Decreto contenente le Regole Tecniche previste anche dall'articolo 22, comma 4³⁷ del suddetto Codice.

Altro motivo, non del tutto secondario, per cui è auspicabile una sollecita emissione del decreto in oggetto è il fatto che l'abrogazione del comma 4 dell'articolo 23 del Codice dell'amministrazione digitale ha fatto sì che attualmente non si faccia più alcun riferimento alla figura del responsabile della conservazione, fondamentale nel processo di dematerializzazione.

Per la verità, formalmente risulta più corretto così in quanto tale figura non era definita prima, né in questa norma né in altre precedenti norme primarie³⁸.

Il decreto in oggetto potrebbe, dunque, porre rimedio anche a questa “continua disattenzione”.

³⁷ Articolo. 22 - Documenti informatici originali e copie. Formazione e conservazione
... Omissis ...

4. Le regole tecniche in materia di formazione e conservazione di documenti informatici delle pubbliche amministrazioni sono definite ai sensi dell'articolo 71, di concerto con il Ministro per i beni e le attività culturali, nonché d'intesa con la Conferenza unificata di cui all'articolo 8 del decreto legislativo 28 agosto 1997, n. 281, e sentito il Garante per la protezione dei dati personali.

³⁸ La figura del responsabile della conservazione, pur essendo molto trattata nella letteratura, è citata solo in circolari e delibere dell'AIPA e del CNIPA, ma mai in norme primarie.