Informatica

Un backup a prova di catastrofe

di Nicola Bortolotti

Di fronte alle calamità naturali, anche le società tecnologicamente avanzate – specie se colposamente dimentiche del proprio passato – si trovano sovente impreparate ad affrontare la conseguente crisi. In tempi recenti, anche culture avvezze alle strategie di "crisis management" si sono fatte cogliere impreparate di fronte all'imponderabile, ad una imprevedibilità riassumibile in uno tsunami di dimensioni superiori alle peggiori previsioni. Nei momenti difficili e tragici, scoprire la sostanziale inefficacia di protocolli di sicurezza spesso vissuti più come sterile e burocratico obbligo di legge che come effettiva e condivisibile esigenza, può sommare ai danni diretti anche cospicui danni indiretti che – in talune tipologie di attività – possono facilmente superare quelli diretti.
Nel caso del recente terremoto emiliano, alcune situazioni di questo tipo sono state direttamente documentate dai telegiornali, sfuggendo all'anelito di minimizzazione e normalizzazione mediatica che ha contraddistinto la copertura di quest'evento: è il caso di uno studio di progettazione che aveva pianificato di traslocare l'attrezzatura proprio il 29 maggio, quando il secondo forte episodio sismico ha fatto crollare l'edificio già precedentemente lesionato provocando la perdita non solo del capitale materiale ma anche di gran parte di quello immateriale, nel caso specifico ben più rilevante.
Su un caso come questo, paradigmatico, si possono fondare numerose riflessioni tra le quali rivestono importanza fondamentale quelle inerenti alle strategie di backup.
È fuori di dubbio, infatti, che lo studio colpito dal disastro effettuasse regolarmente copie di sicurezza, che – tra l'altro – costituiscono anche un obbligo di legge conseguente alla normativa sulla protezione dei dati personali. Con ogni probabilità, tuttavia, tali copie venivano conservate nello stesso edificio collassato e, quasi sicuramente, sono state distrutte dal crollo o sono divenute irrecuperabili.

Il backup altrove

Quali contromisure possono essere prese per evitare situazioni di questo tipo? L'unica via di uscita è quella di custodire almeno un backup in un luogo geograficamente diverso da quello ove è fisicamente ubicato il server, una scelta che presenta – tuttavia – alcuni problemi.
Il primo è di carattere tecnico: dal punto di vista della velocità, infatti, è fuori di dubbio che la scelta più rapida sia quella di effettuare il backup localmente su un numero sufficiente di idonei supporti (che, a seconda dei casi, possono essere nastri magnetici, hard disk, DVD, flash drive...) che verranno poi fisicamente trasportati e custoditi in luogo diverso. Un'alternativa ben più pratica ed elegante sarebbe quella di effettuare il backup su un supporto altrove ubicato (eventualmente anche all'estero) tramite la rete Internet. In quest'ultimo caso, però, bisogna fare i conti con l'estrema lentezza del trasferimento di dati, una velocità che – in certi casi – può essere così esigua rispetto alle esigenze da rendere tale via del tutto impercorribile. Un comune collegamento ADSL, infatti, è caratterizzato da una fortissima asimmetria nelle velocità cosiddette di "download" (ossia scaricamento, cioè trasferimento dal server remoto all'utente) e di "upload" (ossia

caricamento, cioè trasferimento dall'utente al server remoto), con una netta predominanza della velocità di "download" (di norma venti volte superiore a quella di "upload"); tale caratteristica non rappresenta un difetto ma una ottimizzazione poiché – utilizzando il collegamento per "navigare" normalmente – il flusso informativo dai server all'utente è nettamente preponderante rispetto alle richieste dell'utente ai server. Utilizzando l'economica ADSL a fini di backup, tuttavia, la velocità di "upload" diventa il parametro di interesse fondamentale, e la sua strozzatura (tipicamente non supera mai i 500 kbps teorici) porta a velocità di backup decisamente insostenibili, dell'ordine di una sessantina di KB/s (laddove la "B" maiuscola sta per byte). In altri termini, per mezzo di una normale ADSL il contenuto di un semplice CDRom (640 MB) avrebbe bisogno di circa tre ore per essere "backuppato", laddove lo stesso può invece essere scaricato in una decina di minuti nel caso migliore, in condizioni di scarso traffico.

Il secondo problema che sorge, nel caso di backup in "altro luogo" su "altro supporto", è di carattere prettamente legale: come garantire la tutela dei dati personali in luoghi che non siano quelli controllati e controllabili che contraddistinguono l'ambiente di lavoro, quei luoghi dall'accesso sorvegliato che – per rendere meglio l'idea – dovrebbero essere ben definiti e catalogati nel DPS? Questo aspetto è notevolmente sottovalutato non solo in Italia ma in tutto il mondo, in tutti gli ambiti che non siano militari (con gravi lacune anche su quel fronte), con conseguenze che possono andare dal grottesco al pericoloso e che – talvolta – assurgono agli onori della stampa. Si pensi ai danni che potrebbe provocare il semplice smarrimento di una "pen drive" contenente dati riservati, alla dimenticanza o al furto di un intero "notebook", tutti fatti che si verificano con cadenza pressoché quotidiana come testimoniato da numerosi articoli – tra i quali quello reperibile all'indirizzo http://www.ictbusiness.it/cont/news/difesa-inglese-colabrodo-persi-280-computer/27830/1.html – e che così sintetizza la situazione:

"*I dipendenti del Ministero della Difesa di Sua Maestà hanno perso o smarrito 280 computer nel giro di 18 mesi, di cui 188 notebook e 99 desktop. La sbadataggine è arrivata a tal punto da incrementare questo dato già di per sé impressionante, con ulteriori 72 dischi fissi e 73 pendrive USB, che a quanto pare sguizzano fuori dalle tasche con un'agilità da anguille. Non è finita: il conteggio comprende 150 nastri di backup, 18 telefoni cellulari, 10 BlackBerry e 194 CD o DVD*". Una mole di dati impressionante che, se finisse nelle mani sbagliate, sarebbe tale da minare alle fondamenta la sicurezza di qualsiasi nazione, se non che, prosegue sempre l'articolo: "*La magra consolazione è che tutte le informazioni contenute sui dispositivi siano state crittografate o che ci fossero altre misure di sicurezza in atto tali da garantire che nessuno potesse entrarne in possesso, almeno questo è quello che è stato garantito*". Da sottolineare anche che, sempre rimanendo al solo Regno Unito, "*il numero dei prodotti dispersi dagli sbadati agenti di sua Maestà è diminuito rispetto al record negativo toccato nel 2008, quando erano stati persi 326 computer portatili*".

La necessità della crittografia

In realtà, se davvero tutti i dati inglesi erano protetti, la "consolazione" non potrebbe certo dirsi "magra". L'unico modo per trasportare (fisicamente o virtualmente) dati in ragionevole sicurezza al di fuori della propria organizzazione è – infatti – quello di crittografarli con opportuni algoritmi. Esistono numerosi software gratuiti in grado di rendere pressoché inaccessibili ad estranei singoli file o intere partizioni e dischi, ad esempio il consolidato e potentissimo prodotto opensource TrueCrypt (http://www.truecrypt.org), il cui uso è semplicissimo. Alzi la mano, tuttavia, chi non ha mai trasportato al di fuori della propria azienda almeno un file contenente dati personali non adeguatamente protetto. Copiare "in chiaro" è infatti molto più pratico e non costringe ad una metodica come il comunicare la propria password ad altri (o utilizzarne di condivise), che contribuirebbe comunque a diminuire la sicurezza complessiva. Si aggiunga a ciò il fatto di poter visualizzare "al volo" il file su qualsiasi computer, senza dover preventivamente installare alcunché.

Tali obiezioni, tuttavia, non sussistono nel caso di un backup "rimuovibile", che andrebbe sempre e tassativamente effettuato su supporti crittografati con una password aziendale nota all'amministratore e ai responsabili. L'unico problema che potrebbe sorgere è quello legato alla scomparsa dal mercato del software di crittografia (cosa impossibile se si utilizza un programma opensource multipiattaforma come TrueCrypt) o al recupero parziale dei dati nel caso di "crash" del disco o danneggiamento della partizione che – qualora essa sia crittografata – sarebbe effettivamente assai più arduo. Trattandosi, tuttavia, di una copia di backup, il problema non si pone in modo drammatico. Tale crittografia sarebbe auspicabile anche sui file system dei computer portatili, in quanto la protezione dell'accesso mediante password è facilmente aggirabile, così come i "permessi" sui singoli file.

Quanti backup?

Un solo archivio di backup non può mai ritenersi sufficiente, soprattutto nel caso di copie di sicurezza "trasportabili". La ragione è semplice: qualora si verificasse un importante crash (logico o fisico, di qualunque natura) durante il backup, si correrebbe il rischio elevato di perdere sia i dati sia la copia. È bene, quindi, prevedere un minimo di due copie di backup da utilizzare alternativamente (ad esempio con la politica dei giorni "pari" e "dispari"). In tal modo il rischio di perdita sarebbe assai limitato. Nel caso di backup trasportati in luoghi geograficamente differenti, tale strategia è ancora una volta vincente, perché in ogni momento esisterebbe almeno una copia dei dati in un luogo differente da quello principale di lavoro, ma per raggiungere questo scopo è necessario prevederne almeno tre.

Si tenga, infatti, presente il fatto che la copia di sicurezza non può essere fatta su file "in uso" in quel momento e – dunque – viene effettuata di norma nei tempi morti, tipicamente durante la notte; ciò, tuttavia, comporterebbe ulteriori problemi nel caso del trasporto fisico dei supporti di memorizzazione, che potrebbero infatti trovarsi tutti contemporaneamente nella sede dell'azienda e – dunque – rischierebbero di andare simultaneamente distrutti. Per agire in sicurezza, in questo caso, ci sarebbe dunque la necessità di un minimo di tre diversi insiemi di supporti di backup che – per fissare le idee, possono essere denominati A, B e C. Supponiamo che venga portato in sede l'insieme di supporti A (ad esempio una o più cassette o hard disk) e lo si appronti affinché venga sovrascritto nella notte. Il giorno successivo si porterà in sede l'insieme B e lo si predisporrà affinché venga sovrascritto nella notte; il backup A aggiornato verrà portato nel sito "alternativo" di custodia dei backup al termine dell'orario di lavoro. Durante la giornata, quindi, in sede vi saranno entrambi gli insiemi di backup A e B, ambedue vulnerabili; in luogo diverso, tuttavia, vi sarà il set di backup C, preservato da calamità. Al termine della giornata si porterà nel sito di custodia il set A, aggiornato al giorno precedente. Il giorno seguente si porterà in sede il backup C affinché venga sovrascritto durante la notte, mentre – al termine della giornata – si trasferirà l'insieme B nel sito di custodia. L'indomani sarà portato in sede il set A, alla fine si riporterà nel sito di custodia l'insieme C, e il ciclo riprenderà quindi dall'inizio. La strategia è chiara: qualunque cosa possa succedere in sede, in qualunque momento della giornata vi sarà sempre almeno un set di backup integro custodito in un luogo diverso.

Quali politiche di backup?

La strategia di copia di sicurezza – soprattutto se remota o custodita in "altro luogo" – non è irrilevante: totale, differenziale o incrementale? Fermo restando che almeno un backup totale è necessario (all'inizio e a cadenze regolari), che questa copia integrale richiede un tempo elevato per essere effettuata e deve essere presente in tutti i backup "alternativi" (che, come scritto precedentemente, è bene siano almeno due), è raro il caso in cui si decida che tutti i backup (specie se giornalieri) siano totali, in quanto ciò comporterebbe tempi di effettuazione troppo elevati, del tutto insostenibili nel caso remoto.

Tra una copia totale e l'altra (che andrà effettuata con cadenza settimanale o mensile) si può quindi scegliere di effettuare backup giornalieri differenziali o incrementali. Entrambe le opzioni presentano punti di forza e di debolezza.

A favore del backup differenziale (che comprende tutti i file modificati dall'ultimo backup completo) vi è la maggiore robustezza in caso di necessità di "restore", ossia di recupero dei dati, in quanto saranno necessari solo l'ultimo backup completo e l'ultimo backup differenziale; come "contro", vi è il fatto che la dimensione del backup differenziale aumenta sempre, giorno dopo giorno, fino al successivo backup completo. Non è, tuttavia, necessario conservare i backup differenziali precedenti all'ultimo.

Il punto di forza del backup incrementale, invece, è il fatto che comprende solo i file modificati dopo l'ultimo backup totale o incrementale, garantendo le minime dimensioni possibili (fattore importantissimo nel caso di backup remoto via Internet). Il punto di debolezza è che – nel caso in cui sia necessario il "restore" – saranno necessari l'ultimo backup totale e tutti i backup incrementali effettuati dopo di esso, nel corretto ordine. La tecnica incrementale presenta, quindi, maggiore criticità in fase di eventuale "restore", anche se è la più efficiente in fase di backup.

I servizi di "storage" remoto su Internet

Una possibilità aggiuntiva di backup – soprattutto per le piccole realtà – è costituita dagli interessanti servizi di "storage" in rete, un mercato nel quale ha fatto recentemente il suo ingresso anche il colosso Google con il suo "Google Drive". I tre principali fornitori ADrive (www.adrive.com), l'attuale leader Dropbox (www.dropbox.com) e il citato Google Drive (drive.google.com) offrono ad ogni account un numero virtualmente illimitato di gigabyte di

spazio in rete, dietro pagamento di un canone; le offerte gratuite risentono di alcune limitazioni, assai pesanti nel caso di ADrive, che costringe all'utilizzo della sola interfaccia web e rende dunque il servizio "free of charge" inutilizzabile a fini di backup. Nel caso di Dropbox e Google Drive, invece, mediante l'installazione appositi programmi ben interfacciati con il proprio sistema operativo, è possibile far sì che il contenuto di una cartella del proprio computer (e di tutte le sottocartelle in essa contenute) sia copiato in rete (con l'accesso ovviamente protetto da user e password e i file crittografati); il tutto viene silentemente e automaticamente aggiornato ogniqualvolta un file viene modificato o aggiunto.

Il fine di questi servizi non è unicamente quello del backup ma anche quello della condivisione di contenuti tra computer diversi (dello stesso proprietario o gruppo di utenti), un'esigenza divenuta quantomai pressante con il proliferare di strumenti di lavoro diversi (Personal Computer Desktop, Notebook, NetPC, Smartphone) per il quale l'utilizzo della classica "chiavetta" è sempre meno pratico. A ben guardare, però, pur con tutte le riserve del caso (soprattutto sotto il versante della sicurezza e della velocità), nelle piccole realtà di lavoro l'installazione di Dropbox o di Google Drive sulle cartelle condivise del server risolverebbe con il minimo sforzo il problema del backup, addirittura in modalità remota.

Gli archivi digitali sono comunque preferibili

Se le calamità naturali pongono il pressante problema di verificare con grande attenzione l'idoneità delle modalità di backup degli archivi in forma elettronica, va comunque tenuto presente il fatto che un archivio digitale contenente circa cento milioni di pagine cartacee può essere interamente copiato in "locale" in una giornata lavorativa (o nell'intervallo tra una giornata e l'altra). Se si pensa a quanti preziosi archivi cartacei – anche delle amministrazioni pubbliche – sono stati parzialmente ma irrimediabilmente persi in occasione dell'ultimo sisma, ci si può comunque rendere conto di quanto sia vantaggiosa la loro digitalizzazione, anche (e, forse, soprattutto) nel caso di catastrofi. È fondamentale che essi siano conservati in almeno due copie poste in luoghi geograficamente distinti, e questo anche se ci si rivolge a servizi di "storage" su Internet, in quanto persino i migliori "provider" non offrono – di norma – alcuna garanzia in caso di "cause di forza maggiore".

Per quanto concerne la rete Internet, progettata ai tempi della guerra fredda per offrire robustezza di comunicazione anche in caso di attacco nucleare, essa ha confermato in occasione del terremoto la sua straordinaria affidabilità.

Rimane, comunque, una domanda conclusiva non oziosa: se l'edificio dell'azienda della quale siete responsabili collassasse nel fine settimana, i vostri dati sarebbero al sicuro?