

Cos'è un CD-ROM?

Come nasce, come è strutturato e come si fa. Ladies e Gentlemen: il CD-ROM...

Cosa sia un compact disc è cosa ben risaputa, ma come nasce e a quali precise caratteristiche deve aderire forse non è un argomento noto a tutti. In realtà è bene partire dalle specifiche di cui trattiamo: i CD-ROM

Prodromi

Correva l'anno 1980 quando vennero mostrati i primi CD audio, lanciati commercialmente nel 1983, e si decisero i «giochi» per la formattazione e la struttura logica dei dati. Tutto ciò che riguardava i sistemi CD per il formato e le tecniche di registrazione digitale sono stati appannaggio della Philips, mentre il sistema ECC (rilevazione e correzione degli errori) è stato sviluppato congiuntamente dalla Sony e dalla Philips, generando così il famoso «Red Book» contenente tutti gli standard necessari per consentire una totale compatibilità fra i dischi ed i vari tipi di lettori.

La produzione di un CD prevede almeno tre stadi di lavorazione: la pre-masterizzazione, la masterizzazione e la replica dei dischi. Il primo stadio serve a registrare le informazioni audio in forma digitale su nastro, per consentire un input standard per il processo di masterizzazione. Questo consiste nel generare una matrice di vetro usando un fascio laser sulla base delle informazioni presenti sul nastro; dopo ciò la replicazione dei dischi prevede la creazione di vari submaster dal primo e la stampa in grande quantità dei CD veri e propri.

Il materiale base per il pressaggio dei dischi è un tipo di policarbonato detto «macrolon», il quale viene liquefatto e iniettato nelle presse ad alta pressione e temperatura. Dopo questo primo passaggio, in cui viene creata l'unica superficie utile, la parte stampata viene ricoperta da una pellicola di alluminio per produrre lo strato riflettente necessario alla rilevazione dei pit da parte del laser di lettura. La fase finale di produzione del CD è la ricopertura del tutto con la resina plastica, l'etichettamento e, naturalmente, la punzonatura del foro centrale.

Caratteristiche

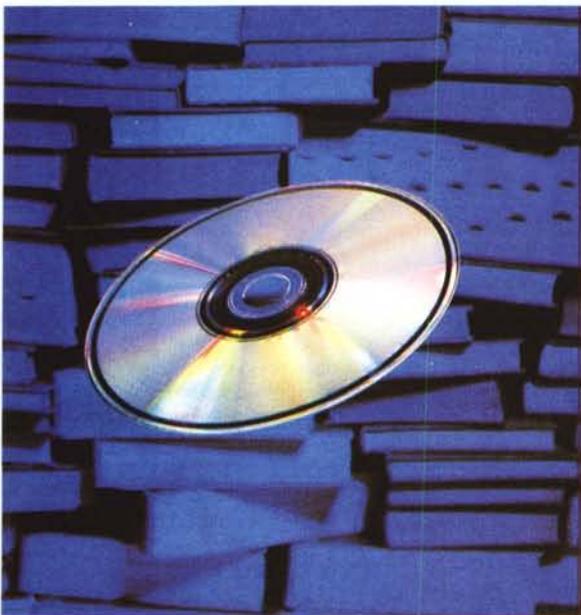
Esistono tre sezioni fondamentali nel CD (di qualunque tipo): quella di start, la parte utente e quella di fine. Ciascuna sezione contiene informazioni codificate in forma digitale: la prima sezione ospi-

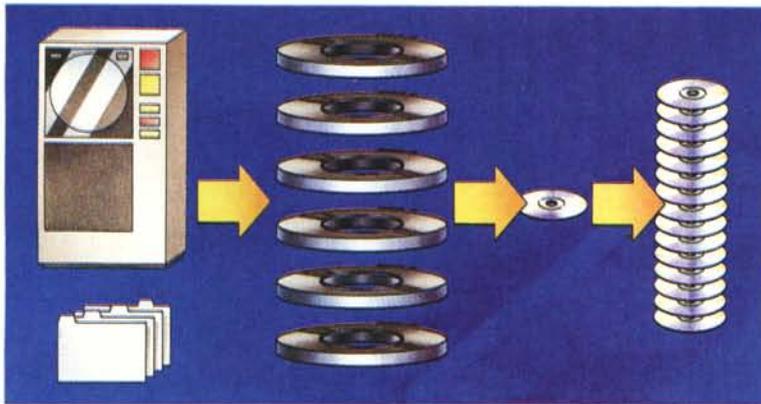
ta la TOC (Table Of Contents) che indica il numero di tracce registrate ed il tempo dall'inizio del disco a ciascuna traccia. L'area di dati utente è autoesplicativa: vi sono tutti i dati che saranno usati da chi ha acquistato il CD: musica, informazioni, immagini eccetera. L'ultima area contiene le informazioni relative alla gestione della fine della riproduzione, ad uso e consumo del player.

I dati nel CD master sono registrati in una spirale che va dall'interno all'esterno, con la possibilità di dividerli in un massimo di 99 tracce. I dischi sono letti con il metodo CLV (Constant Linear Velocity) di 1,25 m/s che garantisce la massima capacità di registrazione, ma impone un accesso più lento rispetto ai sistemi che lavorano a velocità angolare costante (CAV). Infatti il CD ruota ad un numero di giri variabile, da 200 a 430 a seconda dell'area letta.

Sia CD audio che CD-ROM lavorano sui principi sin qui esposti: hanno medesimo formato e composizione chimica, ruotano alle stesse velocità e sono registrati in modo CLV. Adoperano lo stesso sistema di modulazione detto «otto a quattordici» (EFM, Eight to Fourteen Modulation) e il metodo CIRC per la correzione degli errori (Cross Interleaved Reed-Solomon Coding, codificazione Reed-Solomon ad interallacciamento incrociato). Le lievi differenze fra i due tipi di dischi scaturiscono dal fatto che i CD-ROM devono contenere dati e non musica codificata digitalmente; un errore di un bit nella musica è quasi inaudibile, ma può essere catastrofico in una serie di dati: la Philips quando introdusse i CD-ROM fece un divertente esempio, disse che questo bit errato avrebbe potuto modificare un più in un meno in un bilancio di una banca...

Per ovviare a ciò, lo «Yellow Book» dei CD-ROM impone la divisione dei dati in blocchi logici da 2 KByte, ognuno registrato in ciascun settore fisico del CD-ROM. Ogni settore fisico contiene 98 frame di 24 byte, per dare al settore la capacità di 2352 byte, così utilizzati: 12 byte di Sync Field; 4 byte di header; 2048 byte di dati; 288 byte ausiliari.





Questo grafico esemplifica il processo di creazione di un CD-ROM: generazione dei dati, premasterizzazione, masterizzazione, replicazione.

Tutti i dati, tranne i byte di sincronismo sono miscelati (per l'interallacciamento). Al momento gli ultimi 288 byte sono usati nel CD-ROM per un ulteriore livello di correzione d'errore (il cosiddetto Mode 1), mentre nei CD audio vengono registrati altri dati (Mode 2). I due modi possono coesistere sullo stesso media, infatti sono stati prodotti CD-ROM che integrano dati e suono nella stessa applicazione.

Necessità

Dai numeri suesposti appare chiaro che tutta la fase di preparazione dati dovrà collimare con le specifiche necessarie del CD-ROM. Questo significa che se si vuole realizzare un disco per distribuire un database, questo dovrà essere organizzato, indexato e formattato secondo le specifiche CD-ROM, quindi con header, sync ecc.

Se si seguono alla lettera le specifiche del «Yellow Book» per preparare un CD-ROM, questo dovrebbe essere letto virtualmente da qualunque lettore. Ciò è vero, ma il problema è che il CD-ROM contiene dati per computer, quindi è necessario un elaboratore per potervi accedere, e questo implica una serie di problemi di compatibilità.

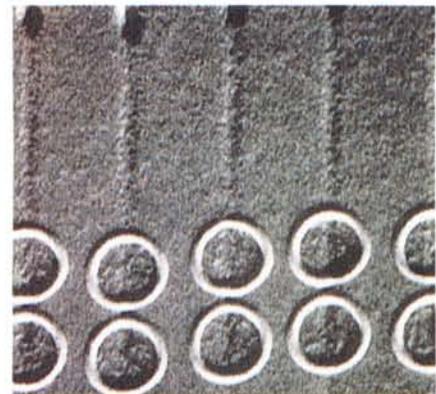
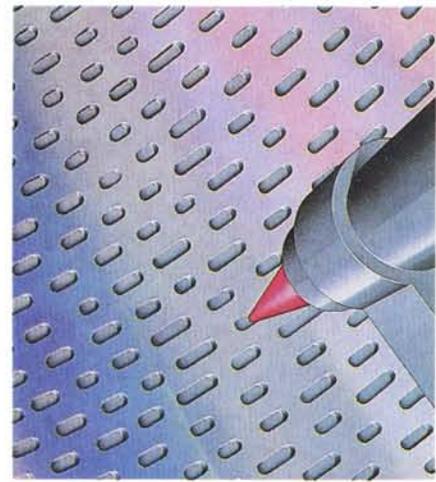
Innanzitutto l'organizzazione logica del CD-ROM è funzione del computer che dovrà leggerlo, in quanto il libro giallo non dà specifiche di questo tipo. In realtà esiste adesso lo standard High Sierra, ma ne parleremo dopo.

Questa dipendenza dell'host determina il tipo di dati e l'organizzazione, ad esempio il testo può essere in EBCDIC anziché in ASCII e adatto ai sistemi 9370 ma illeggibile da Unix.

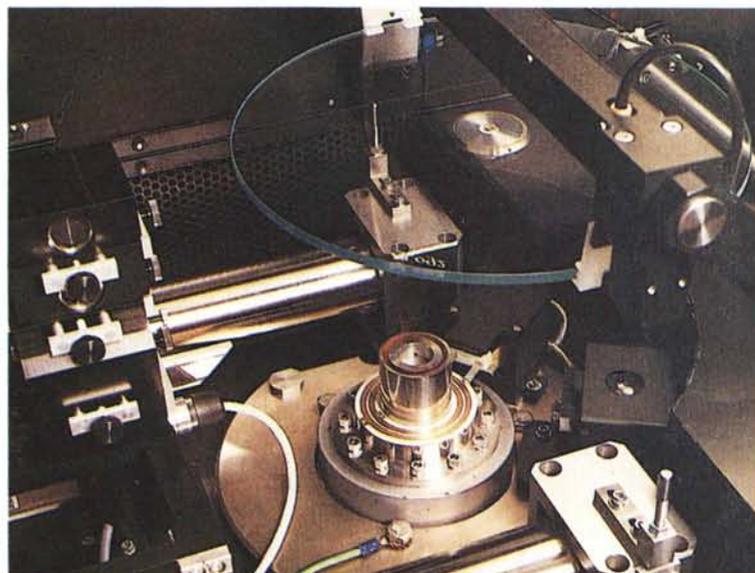
Il secondo punto di discordia è dato dai lettori per CD-ROM. Ogni player ha

un suo controller, perciò ciascun controller ha il suo formato relativo ai comandi che è in grado di accettare. Naturalmente non esiste uno standard fra i costruttori, cosicché è un'impresa far funzionare un software di gestione di un lettore progettato per quelli di marca X con i drive di marca Z. Quindi chi vende software su CD-ROM deve tenere presente che gli utenti possono avere diverse marche di drive, ma il software deve funzionare con tutti.

Un ulteriore problema di compatibilità è dato dal tipo di software di retrieval che si adopera col CD-ROM. Anche in questo caso se realizzo un DB usando per il retrieval BRS (un classico del settore), dovrò formattare tutti i miei



Il CD-ROM è composto da pit (fosses) che alterano la riflessione del laser di lettura.



Questa è la macchina che genera il master in vetro usando il laser (in basso nella foto) gestito dai dati premasterizzati.



Questo è l'interno di un lettore CD-ROM mentre a destra è visibile il relativo controller, spesso fonte di incompatibilità.



dati secondo le necessità del mio programma di consultazione, rendendoli così incompatibili con altri eventuali software che possiedo su altri dischi o su altri sistemi. Per fare un esempio molto semplice, immaginate di avere su CD-ROM 500 MByte di dati di spreadsheet Framework II memorizzati come *. FW2: non riuscirete mai a leggerli usando Multiplan, eppure siete sotto MS-DOS e adoperate lo stesso computer con lo stesso drive per CD-ROM. L'unica soluzione è scrivervi un programma traduttore da un formato di file all'altro, a patto di sapere come sono

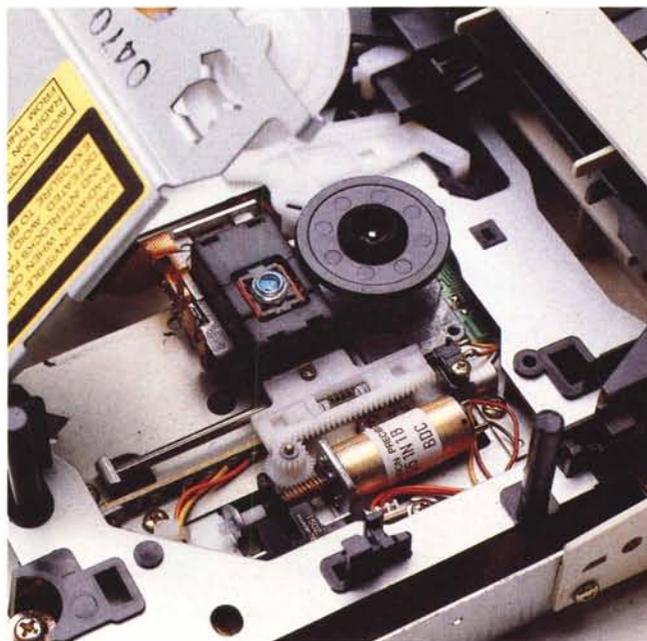
fatti i due file e di saper fare il suddetto programma: probabilmente l'utilizzatore di una serie di dati fiscali all'interno di uno studio notarile è la segretaria, la quale spesso non sa nemmeno cosa sia una struttura o un ciclo do-while...

Organizzazione

Per cercare di ovviare a questi problemi di incompatibilità è nato un gruppo di lavoro formato da varie aziende, con lo scopo di creare uno standard logico per la registrazione dei dati su CD-ROM, quindi per definire i volumi veri e propri.

Il nome dell'associazione è High Sierra Group, ed è diventato anche il nome (HSG) per identificare quei CD-ROM che sono codificati secondo questo standard logico, ormai diventato la norma ISO 9660. L'HSG è nato per i sistemi operativi più comuni: MS-DOS, Xenix e VMS, e prevede l'organizzazione logica dei file a differenti livelli. I file possono essere raccolti in subdirectory, che possono arrivare sino all'ottavo livello secondo la logica di archiviazione dei dati raccolti. Un file è definito nel numero di blocchi ed estensione, con in più la possibilità di avere un eXtended Attribute Record (XAR). Lo standard High Sierra fissa un massimo di 40 file per ogni directory, in modo da velocizzare al massimo le operazioni di accesso ad un file. Infatti non è possibile accedere direttamente al singolo file, ma è necessario eseguire una ricerca lineare su tutto il contenuto della directory. L'indirizzamento alla directory contenente il file da noi cercato è gestito tramite una struttura detta Path Table, che contiene i puntatori alle directory, quindi agli indirizzi di partenza da cui iniziare la ricerca del file voluto.

Questa organizzazione dei dati e le caratteristiche fisiche del CD-ROM, non ne fanno certo una memoria ultraveloce, ma del resto il CD-ROM è nato per essere un supporto affidabilissimo con cui distribuire una quantità spaventosa di dati a costi bassissimi. In questo non lo batte nessuno, ed inoltre un tempo di accesso medio di mezzo secondo, tipico dei drive dell'ultima generazione, non spaventa nessuno, neanche chi è abituato agli hard disk da 28 millisecondi, dato che pochi anni fa lavorava con dischi da 80 ms e oltre...



Particolare della testina di lettura del player, più piccola e con un laser meno potente dei drive WORM.



In-Cat System

Una giovane azienda milanese si ripropone di assaltare il mercato dei CD-ROM: software e sistemi per la produzione sono il suo piatto forte.

Il mercato italiano delle memorie ottiche è abbastanza particolare, le aziende che vi operano normalmente godono di ottima salute grazie a varie attività in altri settori (leggi Hitachi, Philips, Sony, editori vari, fiscali e non). Tutti dicono che la fetta di torta è molto appetitosa, ma al momento è ancora rischioso puntare tutto sul CD-ROM vero e proprio, tranne i realizzatori dei famosi codici tributari, il vero ed unico prodotto ottico italiano di largo consumo.

Il lato della produzione del CD-ROM è stato, almeno sinora, prerogativa di poche aziende o di chi possiede l'unico sistema veramente diffuso per l'in house publishing dei CD-ROM: il CD Publisher della Meridian, una società americana.

La In-Cat System di Milano ha invece introdotto in Italia, la quale è la nazione europea con il maggior numero di CD-ROM, una serie di sistemi e titoli che puntano proprio alla massima diffusione dei dischi ottici, sia come software su CD-ROM, sia come mezzi di realizzazione degli stessi. Su richiesta è in grado di fornire hardware e software in grado di consentire la realizzazione di nastri premasterizzati pronti per essere pressati in CD-ROM, offrendo sistemi con hard-disk da 330 a 2408 MByte a prezzi decisamente concorrenziali rispetto alla Meridian.

Dischi

Al momento la In-Cat distribuisce 5 titoli, per il momento di importazione statunitense. Si tratta essenzialmente di raccolte, sebbene di «roba» differente, il contenuto dei cinque CD è essenzialmente una raccolta. Di che cosa, vi state chiedendo? Semplice di «roba» utilissima.

Il primo titolo è «BLUE SAL SOFTWARE», come dice il nome stesso è una raccolta di programmi. La società

americana Blue Sail distribuisce circa 1000 floppy di public domain software, tutti i programmi dei 1000 dischetti sono quindi raccolti su questo CD-ROM che si rivela una miniera, che viene aggiornata con cadenza quadrimestrale.

ADA CD-ROM è invece un disco di 150 MByte dedicati esclusivamente al linguaggio ADA. Utility per programmatori, tutorial e programmi pronti all'uso sono il contenuto del CD. Tutto il materiale ADA è stato tratto da bbs vari, dal Deposito Federale dei programmi ADA e da vari centri di sviluppo, qualificando così la provenienza del materiale stampato.

Decisamente più particolari sono «PRAVDA ON DISC» e «FEDERAL PROCUREMENT DISC»; il primo è la raccolta di tutti gli articoli pubblicati dalla Pravda, il giornale ufficiale dell'URSS, pubblicati nell'86 e '87. Naturalmente gli articoli sono stati tradotti in inglese, in modo di renderli comprensibili anche all'altro emisfero... Veramente notevole il software di ricerca, che mediante una serie di finestre di dialogo con l'utente consente una rapida individuazione e consultazione degli articoli contenenti le parole o gli argomenti richiesti. Il secondo disco è invece l'archivio di tutti i prodotti acquistati dal governo degli Stati Uniti, i rispettivi fornitori e tutte le modalità di acquisto. Non è certo un CD-ROM che l'utente medio andrà ad acquistare in edicola assieme all'ultimo gioco per l'Amiga: sono dati che servono quasi esclusivamente a quelle aziende che hanno relazioni di affari o fanno import-export con gli USA.

L'ultimo disco commercializzato è «ALDE», ribattezzato «la miniera» dal sottoscritto...

ALDE

Cosa sia ALDE è presto detto: circa 3000 programmi di pubblico dominio e shareware tutti compresi in un solo CD-ROM. Notare ho scritto compresi, non pressati, perché questo CD-ROM uovo (vista la sua «pienezza»...) è stato imbottito ancora più del possibile usando un semplice ma efficacissimo trucco: il

95% dei programmi sono compressi con ARC o con Library o Squeeze, in modo di riunire più file in uno solo e per di più con scarsa occupazione di byte.

Vi sono 15 subdirectory che suddividono gli argomenti principali: comunicazione, utility per DBase, editor e WP, educational, finanziari, giochi, grafica, utilità per Lotus 1-2-3, utility per printer, sorgenti vari, delizie per il programmatore, miscellaneous e due directory relative ai programmi di due bbs: il Loren Jones BBS di Chicago e l'Howard Ekman Terrapin Station di Minneapolis.

Tutto il software è piuttosto recente (fine '87), e ci sono delle delizie da leccarsi i baffi: vi interessano i sorgenti C di ARC? Vorreste mettere le mani su sorgenti GWBASIC del famoso RBBS? Oppure quello che vi serve sono delle routine in C per accedere ai file DB3 e gestirli come vi pare, ovviamente alla velocità del C compilato?

C'è veramente di tutto, dal più scemo giochetto alla space invaders sino al sofisticato compilatore Prolog, un sorgente in C per fare un piccolo BBS sotto UNIX, passando per la lista dei bug del Turbo-C o un potente label maker per floppy, cassette e nastri... Insomma una vera miniera!

Tanto per citare due numeri, vi sono le directory dei programmi messe su due file, che assommata arrivano a 200 KByte, solo di DIR...

Va da sé che non è umanamente possibile provare tutti i programmi presenti per fare una recensione dettagliata, ma vi posso assicurare che ho fatto più di una notte insonne alle prese con questo CD, in esplorazione nei meandri delle directory scoprendo moltissimi programmi utilissimi, ma soprattutto molto funzionali (c'è un Flow Chart Graphic Editor che è una vera libidine...). Naturalmente per poterli usare bisogna prima trasferire i file scelti su disco magnetico, poi de-ARCarli in modo da ripristinare la configurazione di file originari: molto spesso un solo file contiene altri 7-8 file ARCArli che a loro volta si suddividono in 3 o 4 trovandosi così 15 - 20 file partendo da uno solo. È sottinteso che tutti i decompattatori sono inclusi nel disco, altrimenti sarebbero stati dolori per i non possessori di ARC...

L'ultima nota è da fare a riguardo delle tariffe relative all'uso di programmi shareware: sia nell'immane file read.me, sia nella custodia del CD è chiaramente scritto che l'utente è responsabile delle quote associative e delle tariffe relative al software shareware che intende usare.