

Carlo Becchi

Software Hardware

Appunti sulla storia e sul funzionamento teorico
di un personal computer

Versione 1.1c

Licenza – Documento Internet

Questo documento è rilasciato dall'autore per il solo uso personale. E' permessa la stampa, la copia elettronica e la distribuzione purchè in forma completamente gratuita fintanto che il documento rimane integro ed immutato, nel formato e nei contenuti.

Per la distribuzione commerciale e/o organizzata, la pubblicazione completa o parziale all'interno di altri lavori è necessario il permesso scritto dell'autore che può essere richiesto via email scrivendo a Carlo Becchi (deathlord@iol.it).

Qualsiasi altro uso diverso da quelli previsti, perpetuo senza autorizzazione è da considerarsi una violazione del diritto d'autore.

L'ultima versione di questo documento è disponibile al sito dell'autore:

<http://www.deathlord.it/pro/sl>

Introduzione.

Al giorno d'oggi avere un computer in casa non e' un evento raro, tutt'altro! Ciò nonostante ci sono voluti più di cinquant'anni per giungere dalle prime applicazioni belliche alla diffusione attuale. Inoltre il computer non si e' ancora integrato nelle case come gli altri elettrodomestici e rimane, malgrado tutti gli sforzi compiuti in questo campo, un apparecchio di uso e manutenzione relativamente difficile. Ci sono veri e propri computer integrati nei vostri videoregistratori, forni a microonde, hi-fi, televisioni e in decine di altri strumenti di uso comune, ma "IL" computer rimane ancora uno strumento per pochi.

Quando si parla di computer solitamente si sottintende un "***IBM Compatibile***", ossia una macchina basata su processore di marca ***INTEL*** con un sistema operativo di marca ***MICROSOFT***; questo non e' il solo tipo di piattaforma esistente ne' la migliore, ma e' senza ombra di dubbio la più diffusa nel mondo.

Conoscere il funzionamento del computer e' importante per meglio comprenderlo, ma a mio parere e' altrettanto importante capirne la storia e l'evoluzione, perché ormai e' parte della nostra stessa storia e della nostra cultura.

Vediamo come e' nata questa stirpe di computer che da vent'anni, nel bene e nel male, occupa le nostre scrivanie.

La storia cosi' lontana...

Correva la fine degli anni 70 e la piu' grande industria del mondo di calcolatori elettronici era l'***International Business Machines*** meglio nota come ***IBM***. La leadership del mercato la doveva grazie ai ***mainframes***, ossia supercomputer cosi' potenti da poter da soli servire un'intera azienda o universita'. Il mercato pero' stava svoltando verso un'altra direzione, orientata alla costruzione di computer ***piu' piccoli ed economici*** cosi' da permettere alle piccole imprese, se non addirittura alle famiglie di possederne uno.

Fino ad allora IBM aveva sempre costruito le sue macchine da sola, cosi' come i sistemi operativi erano di produzione propria; questa volta pero' alla grande industria mancava tempo e conoscenza per progettare ex-novo un computer completamente differente da quelli solitamente in produzione. La decisione finale porto' all'adozione di un processore e un sistema operativo provenienti dall'esterno.

IBM si rivolse per il microprocessore del suo nuovo computer alla societa' che lo aveva inventato nel 1969: l'***INTEL Corporation***. Infatti questa casa che oggi e' leader nel settore ideo' questo straordinario strumento, attraverso le menti geniali di ***M.F. Hoff*** e ***Federico Faggin***. Fra i modelli disponibili scelse l'***INTEL 8088***, un processore 8/16 bit (cosa questo significhi lo vedremo piu'

tardi).

Era il 1981 e nasceva lo storico **IBM PC**, che con il suo nome fondo' una categoria di macchine: il **PERSONAL COMPUTER**.

Per il sistema operativo fu lasciata al cliente la scelta fra tre diverse soluzioni: il **PC-DOS** di **Microsoft**, il **P-System UCSD Pascal**, e il **CP/M-86** dell'allora leader dei sistemi operativi per piccoli computer, **Digital Research**.

Con un colpo di genio, al limite della veggenza, **Bill Gates** e **Paul Allen**, fondatori di **Microsoft** chiesero a IBM un compenso minimo ed unico, non legato al numero delle copie vendute, riservandosi pero' di vendere il sistema operativo **MS-DOS** (ossia la versione non personalizzata per il PC IBM) ai produttori che avessero voluto immettersi sul mercato con macchine PC-compatibili. Questo era possibile a causa dell'architettura aperta del computer, accessibile percio' anche ad altre aziende senza alcun debito verso IBM.

Un accordo cosi' strutturato porto' IBM a spingere il PC-DOS a causa degli alti guadagni, e per l'utente era comunque il piu' conveniente costando solo 60 \$ contro i 450 dell' P-System e i 145 del CP/M.

Inoltre quasi tutti i produttori di computer e apparecchiature elettroniche si gettarono sul mercato dei PC-Compatibili adottando MS-DOS come sistema operativo e instradando Microsoft sulla via che l'ha portata a diventare la piu' grande azienda di software del mondo.

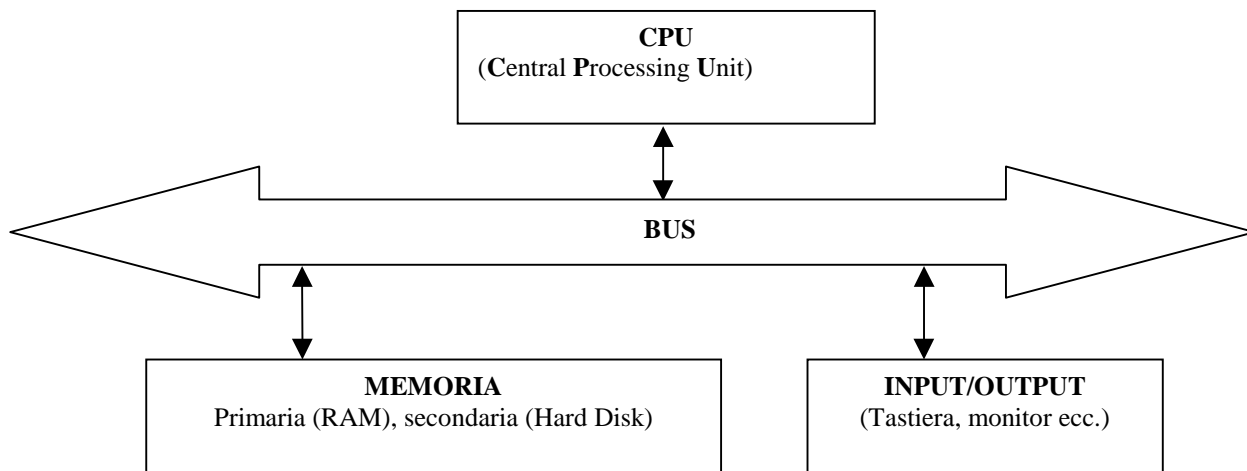
Il resto, come si dice, e' storia.

Funzionamento di un Personal Computer.

In questo capitolo vedremo i principi di funzionamento di un PC, ne definiremo i componenti fondamentali e le unita' di misura indispensabili per confrontare le grandezze in gioco.

Architettura di un PC.

Nonostante i computer siano dispositivi elettronici estremamente complessi, possono essere rappresentati schematicamente come l'insieme di tre "blocchi logici" comunicanti tra di loro: CPU, MEMORIA e I/O.



La CPU, o Central Processing Unit [Unita' di Elaborazione Centrale] e' il vero e proprio "*cervello*" dei computer e su quelli attuali e' rappresentata dal microprocessore; questo blocco si incarica di eseguire tutte le operazioni aritmetiche (conti) e logiche (decisioni, confronti ecc.), ossia fa funzionare i *programmi*.

I dati sui quali la CPU lavora sono immagazzinati nella memoria, che e' presente in due modalita' : *Memoria Primaria* e *Memoria Secondaria*.

La *Memoria Primaria*, detta anche *RAM* (Random Access Memory [Memoria ad Accesso Casuale] o memoria di sistema e' un mezzo di immagazzinamento dati "*allo stato solido*", ossia realizzato mediante circuiti integrati, in maniera simile al microprocessore. Le caratteristiche della RAM sono l'estrema velocita', il fatto di essere sempre interna al computer e di essere *volatile* ossia di cancellarsi quando smette di ricevere l'alimentazione elettrica. Questa memoria e' perciò *temporanea*, e il suo contenuto viene perduto ad ogni spegnimento del computer.

La *Memoria Secondaria*, detta *di massa* permette l'archiviazione permanente di grandissime quantita' di dati su supporti di tipo *magnetico* (hard disk, floppy disk ecc.) o *ottico* (CD-ROM, DVD-ROM). I dati nella memoria secondaria non vengono persi a computer spento, ma il loro accesso e' piu' lento rispetto a quelli nella memoria primaria.

I dispositivi di **I/O** (Input/Output [Ingresso/Uscita]), nonostante siano molteplici possono essere logicamente raggruppati in un unico blocco. Attraverso i dispositivi di Input (tastiera, mouse, joystick, scanner ecc.) l'utente può immettere comandi o i dati da elaborare (ad esempio una lettera...), e il risultato del lavoro svolto dal computer gli viene mostrato mediante l' Output (Monitor, Stampante, Casse acustiche ecc.). Come dovrebbe già essere evidente è questo il blocco che si occupa delle comunicazioni con l'esterno.

I blocchi sono tra loro collegati da un **BUS**, ossia un insieme di fili che trasportano i dati da un'unità logica ad un'altra. Tutti i dispositivi di un computer sono collegati al **BUS**, in modo da rendere la struttura flessibile e suscettibile di future espansioni, in quanto il **BUS** è già predisposto a ricevere nuovi componenti senza bisogno di modifiche fisiche alla circuitazione.

Tutte le parti di un personal computer sono sincronizzate da un orologio interno che scandisce il tempo di esecuzione delle varie istruzioni. Questo orologio si chiama **Clock** e la frequenza dei suoi battiti è misurata in **Hertz** (n. di cicli al secondo), 1000 cicli al secondo sono detti **KiloHertz** (KHz) e un milione di cicli al secondo si chiamano **MegaHertz** (MHz). Assumendo che un processore funzioni a 100 MHz e sia in grado di eseguire un'istruzione per ciclo di clock, allora questo processore compierà 100 milioni di istruzioni al secondo. Per avere un'idea del grande sviluppo dei microprocessori si può notare che il processore dell' IBM PC, l' Intel 8088 aveva un clock, nel 1979 di 4,77 MHz, mentre il più lento dei Pentium III processore del 1999 viaggia a ben 450 MHz! [ATTENZIONE : questo non vuol dire che il Pentium III è 100 volte più veloce di un 8088, in quanto questa affermazione sarebbe vera solo nel caso in cui veramente un processore impiegasse un ciclo di clock per ogni istruzione. L' 8088 poteva impiegare anche 10-20 cicli per eseguire una sola istruzione mentre in particolari condizioni il Pentium III può eseguire più di un'istruzione per ciclo di clock, rendendolo migliaia di volte più potente del suo antico progenitore].

Funzionamento di un PC.

Vediamo adesso qual'è il tipico funzionamento di un Personal Computer.

Per prima cosa l'utente deve avviare il programma, ossia la serie di istruzioni che la CPU deve compiere, che solitamente si trova immagazzinato nella memoria di massa (es. un CD-ROM). Con questa operazione, il programma viene copiato dalla memoria secondaria a quella primaria, all'interno della quale verrà eseguito, istruzione per istruzione, copiando dati in piccole zone di memoria interna al microprocessore chiamate registri. A questo punto, la CPU preleva i dati da elaborare (dalla memoria di massa o da un dispositivo di Input), li elabora, e mette il risultato nella memoria primaria, per una eventuale visualizzazione su di un dispositivo di uscita. Se non si vuole perdere questo risultato sarà necessario "salvarlo" su di un dispositivo di memoria secondaria (Es.

un dischetto).

Tutti i programmi, dalla complessa simulazione scientifica al videogioco adottano questo principio di funzionamento.

Sistema operativo.

Un ***Sistema Operativo*** (SO) e' un programma di base che fornisce le principali funzioni per l'utilizzo delle componenti elettroniche che costituiscono il computer. Senza un SO il computer non puo' funzionare in quanto e' questo software a gestire la memoria, sia primaria che secondaria e le comunicazioni fra i vari componenti che lo costituiscono. I sistemi operativi moderni come Windows 95, 98, NT e 2000 ***costituiscono uno strato di astrazione fra la macchina e i programmi.*** In questo modo non e' necessario per chi scrive un'applicazione tener conto delle schede montate sulla macchina, visto che avra' a disposizione una serie di funzioni standard; sara' compito del produttore della scheda fornire un programma chiamato ***DRIVER*** che fa eseguire al suo prodotto specifico queste funzioni.

Unita' di misura.

Anche i muri sanno che i computer adottano un tipo di codifica detta binaria, ossia composta esclusivamente da ***1*** e ***0***. Si utilizza il codice binario per rappresentare qualunque cosa: numeri, lettere, immagini, suoni e le stesse istruzioni che la CPU deve eseguire.

L'unita' binaria e' percio' quell'entita' che puo' assumere due stati o valori (0 e 1) ed e' chiamata ***BIT***, contrazione di ***Binary digiT*** [Cifra binaria]. Con un bit non sono molte le cose reali che si possono rappresentare: un numero da 0 a 1, una lampadina accesa o spenta... percio' si e' deciso di raggruppare ***otto bit*** e chiamare questo nuovo "oggetto" ***BYTE***.

Il calcolo combinatorio ci informa che se abbiamo otto unita' che possono assumere due stati distinti otterremo $2^8 = 256$ ***combinazioni diverse*** ossia i 256 numeri binari compresi tra ***00000000*** e ***11111111***. Con un byte e' possibile rappresentare ad esempio un numero compreso tra 0 e 255, oppure tutti i caratteri dell'alfabeto sia maiuscoli che minuscoli piu' la punteggiatura e le lettere nazionalizzate o ancora assegnare un colore ad ognuna delle combinazioni ottenendo cosi' 256 sfumature diverse.

Bit e byte raramente si utilizzano da soli: se un bit rappresenta l'accensione o lo spegnimento di un punto sullo schermo del computer (***pixel***) un gruppo di bit con questa funzione rappresentera' un'immagine in bianco e nero. Se allo stesso scopo utilizziamo un byte, l'immagine non sara' piu' composta da pixel che possono essere solo accesi o spenti (ossia bianchi o neri) ma da punti che

possono assumere 256 valori, ossia 256 colori differenti. Quando si parla di "*immagine a 8 bit*" si intende che ogni pixel e' codificato da 8 bit ossia un byte e percio' sara' un'immagine che puo' contenere al massimo **256 colori**. Allo stesso modo immagini a **16 e 24 bit** contengono un massimo di **65.536** e **16.777.216** colori rispettivamente. Un'immagine a 24 bit e' detta **TRUE COLOR**, in quanto contiene un numero maggiore di colori di quanti non possa distinguerne l'occhio umano. Visto che il byte e' un'unita' di misura molto piccola, spesso si utilizzano i suoi multipli:

Kilobyte (Kb) = 1024 byte

Megabyte (Mb) = 1024 Kb = 1.048.576 byte

Gigabyte (Gb) = 1024 Mb = 1.073.741.824 byte

Terabyte (Tb) = 1024 Gb = beh, ci siamo capiti!

Visto che si e' tentato di uniformare questi multipli binari ai multipli decimali definendo 1 Kb = 1000 byte, snaturando la natura stessa (binaria) delle grandezze in gioco, chi ne trova vantaggio considera appunto i multipli binari: ad esempio i produttori di hard disk.

Se vi viene venduto un disco da 10 Gigabyte, si considerano 10.000.000.000 byte, ossia 9,31 Gb binari. Non stupitevi percio' se il sistema operativo rileva un hard disk piu' piccolo di quanto c'e' scritto in etichetta.

Breve lista dei processori Intel utilizzati nei PC

- **8086 (1978):** Processore a 16 bit utilizzato nell'IBM XT, successore del PC originale
- **8088 (1979):** Il processore usato sul PC era internamente un processore a 16 bit, ma il bus dati era a 8 bit, in modo da risparmiare sulla circuitazione esterna. In effetti il PC aveva 8 slot di espansione ad 8 bit, visto che l' 8088 comunicava con l'esterno solo con questa ampiezza. Ovviamente per trasmettere dati a 16 bit era necessario il doppio del tempo.
- **80286 (1982):** Processore a 16 bit con bus per gli indirizzi a 24 bit, che gli permetteva di utilizzare un massimo di 16 Mb di memoria, una quantita' enorme per l'epoca. Questo processore disponeva di un modo reale (real mode) attraverso il quale funzionava come un 8086 veloce, ma con le stesse limitazioni e un modo protetto (Protected Mode) che gli permetteva di indirizzare tutti i 16 Mb e di accedere a servizi per il multitasking, ossia per poter far girare piu' programmi contemporaneamente. Il modo protetto aveva pero' il grande difetto che una volta attivato non si poteva tornare alla modalita' reale di partenza se non con un soft reset del microprocessore. Fu montato per la prima volta sull' IBM AT che evolveva a 16 bit l'architettura del PC. Venne introdotta la memoria virtuale, ossia la possibilita' di indirizzare piu' memoria della RAM effettivamente presente sulla macchina (ad esempio su disco). Il 286 poteva indirizzare 1 GB di memoria virtuale
- **80386 (1985):** Primo processore Intel a 32 bit, sia come registri, bus dati e bus indirizzi (4 Gb di memoria indirizzabile). Oltre al modo reale e protetto a 32 bit disponeva di un modo chiamato Virtual 8086 che rendeva possibile la simulazione della presenza di piu' processori 8086 permettendo a programmi adeguati di fare girare piu' applicazioni in modalita' reale. Il primo computer a montarlo non fu IBM ma Compaq, una prova significativa dell'imminente perdita della leadership di IBM. Nonostante il grande potenziale questo processore fu sfruttato pochissimo, in quanto per molto tempo non ci furono applicazioni e sistemi operativi adatti ad utilizzarne la modalita' protetta a 32 bit. Questo fu il primo processore ad essere accoppiato ad una memoria cache, ossia una memoria statica molto piccola, ma velocissima (10 volte piu' della primaria) posta tra il processore e la RAM di sistema allo scopo di velocizzare l'accesso a quest'ultima; in seguito questa tecnologia fu applicata anche ai 80286 ancora in produzione. Fu prodotta una versione chiamata SX con il bus dati a 16bit. Entrambe le versioni potevano indirizzare 64 terabyte di memoria virtuale.
- **80486 (1989):** Essenzialmente un 386 turbo con un coprocessore matematico incorporato e una cache di 8kb interna. Il coprocessore matematico era sempre stato disponibile come opzione gia' dal tempo dell'8086 per quegli utenti che necessitavano di calcoli in virgola

mobile (decimali con notazione mantissa-esponente). Con il 486DX coprocessore e cache sono presenti all'interno del microchip; questa cache viene chiamata "di primo livello" mentre quella esterna, ancora presente, diventa di "secondo livello". Furono prodotte le varianti SX senza coprocessore matematico, DX/2 con la velocita' di clock interna doppia di quella esterna, DX/4 con la velocita' interna triplicata, una maggiore quantita' di cache di primo livello (16 kb) ed un minor voltaggio di alimentazione.

- **Pentium (1993):** La struttura interna viene stravolta rispetto al 486, con l'introduzione dell'architettura superscalare, di una cache di primo livello di 8 Kb per i dati, di una da 8 Kb per le istruzioni e del parallelismo interno per l'unita' degli interi. Avendo due unita' che lavorano contemporaneamente a 32 bit, il bus dati e' stato portato a 64 bit. Essendo registri e bus indirizzi a 32 bit la quantita' di memoria disponibile e' la stessa che per il 386. Per la prima volta viene introdotta la predizione dei salti che permette di velocizzare questo tipo di istruzioni.

Intel abbandona il nome numerico (il Pentium sarebbe stato il 586) a causa dell'impossibilita' di registrare i numeri come marchi di fabbrica, e per impedire ad altre ditte (AMD e Cyrix) di costruire processori compatibili con lo stesso nome, come era accaduto per 286, 386 e 486.

- **Pentium Pro (1995):** Processore di sesta generazione nato ed esclusivamente utilizzato per i server, ossia i grandi computer che gestiscono reti di calcolatori. Un'architettura superscalare ancora piu' spinta, un coprocessore matematico potente con doppia unita' a virgola mobile, l'esecuzione dinamica delle istruzioni e una cache sincrona (stessa velocita' di clock del processore) di secondo livello da 512 Kb o da 1 Mb integrata sul chip lo hanno reso per molto tempo il re incontrastato di server e workstation avanzate. La sua struttura pero' non era adatta ad utilizzare codice a 16 bit, e percio' nemmeno Windows 95 che ne contiene ancora una buona parte. Su questo sistema operativo il Pentium Pro raggiungeva le prestazioni di un comune Pentium a parita' di clock, ma costava fino a quattro volte tanto!
- **Pentium MMX (1996):** Nonostante sia questa una versione potenziata del comune Pentium, con 32 Kb di cache di primo livello, questo processore merita particolare menzione perche' rappresenta la prima sostanziale modifica alle istruzioni dei processori Intel dai tempi del 386. Vengono infatti integrate 57 istruzioni SIMD chiamate MMX, per l'accelerazione di applicazioni multimediali come giochi, audio, video e riconoscimento vocale. Troverete maggiori informazioni nella sessione "microprocessori" del prossimo capitolo "Nuove tecnologie".

- **Pentium II (1997):** Essenzialmente un Pentium PRO ottimizzato anche per le istruzioni a 16bit, con tecnologia MMX, senza cache di secondo livello (512k) integrata sul chip, ma comunque incorporata nella stessa struttura. La cache non e' piu' sincrona ma viaggia alla meta' della velocita' di clock del processore. Processore e cache sono racchiusi in una cartuccia di plastica detta **SECC** (Single Edge Contact Cartridge). Per la prima volta viene abbandonato il Socket a vantaggio dello slot per il montaggio della CPU sulla scheda madre. Disponibile con bus a 66 e 100 Mhz e in versione Xeon con cache di secondo livello esterna ma sincrona nei tagli da 512k, 1Mb e 2Mb per l'uso nei server.
- **Celeron (1998):** Versione economica del Pentium II, inizialmente progettato per non avere cache di secondo livello, con la versione A (chiamata anche Mendocino) integra sul chip 128 k di cache sincrona, che gli permette di avere prestazioni pari ad un Pentium II di uguale clock. Per ora disponibile alla velocita' di bus 66 Mhz, sia per slot 1 che per socket (PGA).
- **Pentium III (1999):** Poche le innovazioni apportate per ora dal Pentium III, se non un controverso numero di identificazione del processore e un nuovo set di istruzioni SIMD chiamate SSE o KNI che serviranno per accelerare giochi e applicazioni tridimensionali e per facilitare la riproduzione di filmati MPEG-2 (quelli dei DVD). La novita' piu' eclatante e' l'aggiunta di una nuova modalita' operativa (cosa che non succedeva dai tempi del 386) che permette di eseguire contemporaneamente istruzioni SSE a virgola mobile, istruzioni del coprocessore matematico (x87) a doppia precisione ed infine istruzioni MMX. Quando uscirà il relativo chipset Camino (Intel 820) allora verranno introdotte altre sostanziali novita' come l' AGP 4X, frequenze di bus maggiori, un nuovo tipo di memoria dalla banda passante spaventosa (RamBus), nuove specifiche per l'USB [Vedi "Nuove Tecnologie"]. Come il Pentium II anche questo processore e' disponibile in versione Xeon per l'utilizzo nei server..

Nuove Tecnologie.

In questo capitolo affronteremo tutto cio' che offre di nuovo il mercato dell'hardware. Le definizioni dei termini piu' tecnici sono disponi nella sessione "Glossary".

Acceleratori Hardware.

I microprocessori hanno sempre avuto come punto di forza la versatilita': infatti sono dispositivi programmabili che possono compiere qualunque tipo di funzione partendo da un set di istruzioni di base impresso nel loro *microcodice* (le istruzioni direttamente presenti in hardware). Cio' pero' comporta una svantaggio dal punto di vista delle prestazioni, visto che i compiti vengono svolti sempre da programmi, che scompongono il problema in procedure sempre piu' semplici finche' e' possibile affrontarlo in termini binari attraverso le istruzioni del microprocessore. Questi passaggi impiegano tempo, ed e' per questa ragione che si e' pensato di utilizzare per compiti altamente specifici circuiti appositamente progettati che disponessero di istruzioni hardware in grado di svolgere funzioni complesse. L'avvento di Windows, ai tempi una pesante interfaccia grafica per MS-DOS, decreto' che il primo campo nel quale intervenire dovesse essere quello della grafica bidimensionale per le *interfacce utente* (Graphic User Interface o GUI). Un esempio di questo tipo di accelerazione e' il tracciamento di figure geometriche bidimensionali: senza l'hardware dedicato se il processore deve disegnare un quadrato (ad esempio lo sfondo di una finestra di Windows) deve calcolare punto per punto la superficie del quadrato e copiarla nella memoria video. Con l'acceleratore la CPU demanda il compito alla scheda grafica indicandole solo di disegnare un quadrato dal punto (x1,y1) al punto (x2,y2) di colore z, e intanto puo' dedicarsi ad altri compiti mentre la scheda grafica esegue l'ordine assegnato (molto velocemente visto che e' progettata appositamente).

Di recente si e' visto il boom degli acceleratori per la grafica tridimensionale, che in questo caso non aumentano soltanto la velocita' di esecuzione dei programmi (nella fattispecie giochi), ma permettono una qualita' grafica ed effetti speciali altrimenti irrealizzabili in tempo reale.

In ordine di tempo, invece, l'ultima novita' in questo campo sono gli acceleratori per audio posizionale tridimensionale, che permettono di disporre di effetti audio dinamici e avvolgenti anche con due soli altoparlanti.

La grafica 3D e le nuove istruzioni per i microprocessori.

Con l'uscita del processore Pentium Intel inizio' una politica di accentramento, mediante la quale voleva convincere che il solo processore era necessario a svolgere tutte le funzioni di un computer, compresa l'elaborazione audio e video. In effetti il notevole salto prestazionale tra il 486 ed il

Pentium sembrava dare credito a questa ipotesi. Inoltre il mercato sembrava rivolto sempre di più ai giochi tridimensionali che per tradizione, sono molto esigenti per quanto riguarda le prestazioni del processore. Intel allora introdusse per i suoi processori Pentium la tecnologia MMX, ossia un nuovo set di istruzioni di tipo SIMD funzionanti sui numeri interi. SIMD vuol dire Single Instruction Multiple Data, ossia istruzione singola per dati molteplici; in questo modo è possibile applicare un'istruzione ad un gruppo di dati piuttosto che ad uno solo. Questo tipo di operazione si esegue spesso quando si ha a che fare con grafica tridimensionale e riproduzione audio e video. MMX svolge queste operazioni in hardware, perciò molto velocemente.

Il fatto che le istruzioni MMX operino solo sui numeri interi le mise in concorrenza con le schede acceleratrici 3D, e non potendo tenere il passo di queste ultime ne decretò l'insuccesso. Prima di andare avanti però è necessario capire come funziona la grafica tridimensionale e che tipo di grandezze entrano in gioco nei vari stadi della realizzazione di una scena 3D.

La grafica 3D

La composizione di un'immagine tridimensionale su di uno schermo bidimensionale si ottiene mediante quattro passaggi fondamentali:

Tessellation e Physics Phase: in questa prima fase viene definita la fisica degli oggetti che saranno rappresentati sullo schermo e vengono suddivisi in poligoni elementari (solitamente triangoli). In effetti ogni componente di una schermata grafica è composto da un numero elevatissimo di triangoli: più sono e maggiore sarà il realismo dell'oggetto. *Questa fase è realizzata mediante l'uso di calcoli in virgola mobile.*

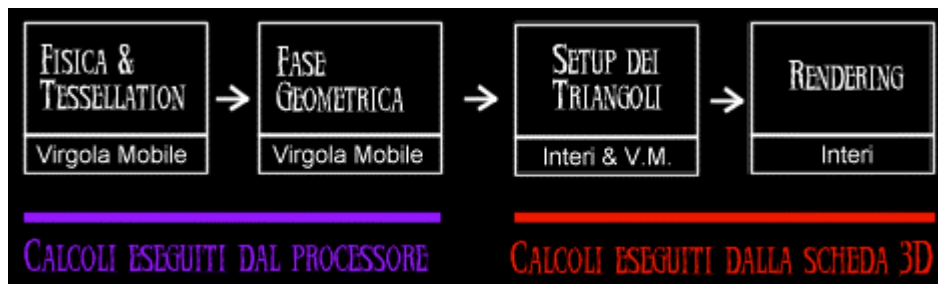
Traformazione Geometrica, Clipping, Illuminazione e Proiezione: a questo punto si trasformano i modelli matematici ottenuti nella Physics Phase per creare rappresentazioni visive degli oggetti. In questa fase gli oggetti vengono posizionati e ridimensionati nello spazio così come viene richiesto dal programma (Trasformazione Geometrica). Attraverso il Clipping vengono eliminate le aree presenti al di fuori dello schermo, poi viene calcolata l'illuminazione degli oggetti e la scena tridimensionale così ottenuta viene proiettata sul piano dello schermo in modo da ottenere "l'ossatura" dell'immagine da rappresentare. *Anche questa fase fa un uso intensivo dei calcoli in virgola mobile.*

Setup dei Triangoli: nella penultima fase i triangoli che compongono l'immagine vengono preparati alla visualizzazione e viene calcolata la variazione dell'intensità del colore tra i vertici in modo da poter modificare le texture nell'ultima fase. *Il setup dei triangoli utilizza sia operazioni sugli interi che operazioni in virgola mobile.*

Rendering: l'ultima fase e' senza dubbio una delle piu' importanti, infatti agli oggetti rappresentati come poligoni complessi viene applicata un'immagine bidimensionale, chiamata texture, che ne rappresenta l'aspetto superficiale, vengono eliminate le facce non visibili degli oggetti, vengono applicati effetti speciali come la nebbia, i riflessi, le ombre, le trasparenze, i filtri che impediscono ad un'immagine ingrandita di risultare sgranata ecc. A questo punto con la rasterizzazione l'immagine rappresentata matematicamente viene convertita in punti (pixel) e diventa pronta per la rappresentazione sullo schermo.

Quest'ultima fase, che comprende esclusivamente calcoli sugli interi, non e' complessa come le prime tre, ma richiede operazioni semplici per un numero elevatissimo di punti.

Nelle attuali configurazioni i primi due passi vengono svolti dal microprocessore, ed in particolar modo dall'unita' per i calcoli in virgola mobile (FPU), mentre il setup dei triangoli ed il rendering sono a cura delle schede grafiche con accelerazione 3D.



Come dovrebbe ormai essere chiaro, essendo le istruzioni MMX in grado di accelerare le operazioni sui numeri interi, sono state poste in diretta concorrenza con le schede acceleratrici 3D, ben piu' performanti e in grado di restituire una qualita' di visualizzazione assai superiore. Per questo la politica di Intel falli': nessun set di istruzioni sara' mai piu' veloce dell'hardware dedicato, invece da una proficua collaborazione possono nascere prestazioni esaltanti. Perche' cercare di contrastare le schede grafiche 3D che rappresentano l'eccellenza nel campo del rendering, e in quelle piu' recenti anche del triangle setup, quando c'e' ancora cosi' tanto da lavorare nelle prime due fasi della grafica 3D a carico del processore?

Fu questa nuova presa di posizione che guido' la nascita di una nuova generazione di SIMD.

Le nuove istruzioni

Attualmente esistono tre tipi di istruzioni SIMD: MMX di Intel, presenti ormai su tutti i processori in commercio e che lavorano sui numeri interi; 3DNow! di AMD, presenti solo sui processori K6-2 e successivi che lavorano sui numeri in virgola mobile ed infine le nuove SSE di Intel presenti sui

Pentium III, ancora per i calcoli in virgola mobile. Mentre i processori con 3DNow! e SSE hanno anche le istruzioni MMX, i due set per i calcoli floating point non sono compatibili tra loro.

MMX: Introdotta da Intel con il processore Pentium MMX questo set di 57 istruzioni permette di applicare un'istruzione ad un gruppo di numeri interi per avere il risultato contemporaneamente per tutti i numeri, con un notevole risparmio di tempo. Questo accadeva perché se prima era necessario applicare ad ogni dato un'operazione, ora se gruppi di dati devono essere sottoposti alla stessa operazione possono essere impacchettati ed elaborati con la relativa istruzione MMX, ottenendo più risultati nello stesso tempo in cui, senza MMX se ne otteneva uno solo. Ci sono però dei problemi: per prima cosa, come già affermato, lavorando sugli interi le istruzioni MMX si "sovrappongono" all'operato delle schede acceleratrici, ben più performanti e piuttosto economiche; più tecnico invece il secondo intoppo. Per immagazzinare i dati per le istruzioni MMX all'interno del processore, vengono utilizzati i registri riservati dell'unità a virgola mobile. Questo fatto impedisce l'uso contemporaneo di calcoli floating point e MMX ed introduce delle latenze nel passaggio tra le due modalità. Inoltre ormai sappiamo che queste operazioni si susseguono di continuo nella generazione di immagini tridimensionali e perciò un approccio di questo tipo non risulta efficiente. L'uso di acceleratori 3D e l'integrazione di funzioni video nelle schede grafiche hanno limitato moltissimo l'uso di MMX. Un ottimo esempio della relativa validità di queste istruzioni è la versione software (non accelerata) del videogioco UNREAL, che riesce a proporre effetti solitamente disponibili esclusivamente con l'ausilio di acceleratori hardware 3D (ed in particolare il filtro bilineare, la nebbia volumetrica ed i riflessi di luce.)

3DNow!: Queste istruzioni di AMD sono le prime SIMD a virgola mobile su un processore di tipo x86 e non furono implementate sui processori K6 per ottenere prestazioni rivoluzionarie, ma piuttosto per rimediare alle performance davvero povere dell'unità a virgola mobile di queste CPU rispetto a quelle dei concorrenti Pentium e Pentium II. Lo scopo che AMD si era prefissata è stato in parte raggiunto, nel senso che le 21 nuove istruzioni per dati a virgola mobile a singola precisione, in effetti, hanno permesso al K6-2 di raggiungere, se non di superare in qualche caso le prestazioni dei Pentium II di pari frequenza di Clock, ma d'altro canto tutte le istruzioni SIMD richiedono software appositamente realizzato, ed il supporto al set di AMD non è stato totale. Tecnologicamente parlando, possono essere eseguite fino a due istruzioni 3DNow! per ciclo di clock che lavorano su registri a 64 bit che contengono ognuno due dati da 32 bit. Perciò si ottiene una velocità di picco di 4 istruzioni eseguite per ciclo di clock. Ancora una volta, come per il set MMX, non è possibile utilizzare contemporaneamente istruzioni 3DNow!, MMX, ed il

coprocessore matematico (x87); in questo caso però il problema, per computer ben equipaggiati è minimo. Infatti non ha senso usare contemporaneamente il coprocessore e 3DNow!, in quanto entrambi lavorano sui dati in virgola mobile, mentre l'inutilità delle istruzioni MMX diventa palese in presenza di un acceleratore grafico. Si noti che parlando di istruzioni in virgola mobile ci si riferisce a quel sottoinsieme necessario per la realizzazione della grafica tridimensionale, e non a tutte quelle disponibili su un processore attuale. L'uso di queste istruzioni è supportato a partire dalla versione 6.0 delle librerie di Microsoft Direct X.

SSE: Le Streaming SIMD Extensions sono le più nuove e complete istruzioni SIMD, presentate per la prima volta da Intel sul processore Pentium III. Prima della presentazione ufficiale di questo microprocessore erano note come Katmai New Instructions (KNI), dal nome in codice del Pentium III, Katmai appunto. L'aggiunta di una nuova modalità operativa (le ultime sono state il V8086 e il modo protetto a 32 bit del 80386 nel 1985) elimina i problemi di MMX e 3DNow!, permettendo al Pentium III di eseguire contemporaneamente istruzioni SSE, x87 classiche a singola e doppia precisione, e MMX. Nonostante sia possibile eseguire una sola SSE per volta (al contrario il K6-2/3 esegue 2 3DNow! per ciclo di clock) l'introduzione di otto nuovi registri a 128 bit (che contengono perciò 4 dati a 32 bit l'uno) permettono l'applicazione di una singola istruzione a 4 dati contemporaneamente, ottenendo prestazioni di picco di 4 elaborazioni al secondo, come per 3DNow!.

Le nuove istruzioni, in totale 70, sono così suddivise: 50 sono le istruzioni SIMD vere e proprie per i calcoli in virgola mobile a singola precisione, 12 servono come estensione per il set MMX in modo da accelerare le prestazioni video MPEG-2 (DVD), e lavorano sui numeri interi, infine le ultime 8, le più innovative consentono al software il controllo diretto il flusso dei dati tra la memoria principale, la cache ed il microprocessore. Mediante Direct X 6.1 è già possibile sfruttare queste nuove istruzioni.

Sito Internet:

Le informazioni riportate in questo testo sono tratte dal sito Internet *DeathLord' Hardware*

Cathedral:



Ideato e realizzato da Carlo Becchi.

Sul sito, raggiungibile all'indirizzo <http://www.deathlord.it/hardware/> questo documento è costantemente aggiornato con le ultime novità del mercato.

Per informazioni scrivere a deathlord@iol.it

Supportate



[segnalando errori ed imperfezioni a deathlord@iol.it](mailto:deathlord@iol.it)

First Version: 14 Luglio 1999. Last Major Update: 12 Agosto 1999 Last Lesser Update: 4 Maggio 2001
--