

Intel Pentium Pro

di Andrea de Prisco



Del Pentium Pro (già conosciuto in passato sotto il nome in codice di Intel P6) ne abbiamo già parlato sulle pagine di MCmicrocomputer in più di un'occasione. Si tratta (per quei pochi che ancora non lo sapessero) del successore del Pentium e rappresenta la sesta generazione dell'architettura Intel. Una delle caratteristiche fisiche del nuovo microprocessore "tecnologicamente più appariscenti" riguarda il fatto che non si tratta semplicemente (si fa per dire...) di un singolo "pezzo" di silicio, ma di una affiatatissima coppia. Per la prima volta nella storia dei microprocessori Intel è stata inserita nel medesimo contenitore ceramico del microprocessore anche la cache di secondo livello (in tagli da 256 o 512 KB), che comunica con il "chip" vero e proprio attraverso un bus interno ottimizzato ad alta velocità (a 64 bit), ben

al riparo da ogni possibile collo di bottiglia dovuto all'utilizzo di qualsivoglia logica esterna.

Se, come è giusto che sia, lasciamo mentalmente da parte la doppia cache integrata e ci concentriamo maggiormente sugli aspetti della progettazione logica del microprocessore vero e proprio (che incorpora, tra l'altro, ben 5.5 milioni di transistor... tutti pronti a farci saltare nell'iperspazio) possiamo apprezzare l'alta sofisticazione tecnologica della "Esecuzione Dinamica" delle istruzioni. È questa, senza ombra di dubbio la "trovata" (anche qui... si fa per dire) più interessante del nuovo nato: una combinazione di tecnologie comprendenti la predizione multipla di salto, l'analisi del flusso dati e l'esecuzione speculativa che permette alle unità di elaborazione interne di venir rifornite ininterrottamente di dati e istru-

zioni da eseguire. Con l'analisi del flusso di istruzioni effettuato durante l'esecuzione del codice il Pentium Pro effettua per suo conto un riordinamento delle istruzioni da eseguire che è indipendente dall'ordine stabilito del programma, ma che ne ottimizza l'esecuzione complessiva. Ogni istruzione - ricordando in questo le architetture data flow - viene eseguita non appena sono disponibili i dati sui quali operare, indipendentemente (se non esistono dipendenze logiche) dall'esecuzione delle altre istruzioni. Questo permette di sfruttare al massimo le capacità di calcolo del processore, ottimizzando in tempo reale il codice non ottimizzato per l'architettura Pentium Pro, senza mai lasciare inoperativa l'unità di esecuzione anche quando si è in attesa di dati e istruzioni non disponibili né nella cache di primo livello né in quella esterna/interna.

Un'architettura data-flow

Il nuovo approccio proposto da Intel per il suo – rivoluzionario – Pentium Pro elimina la restrizione della sequenzialità delle istruzioni tra le tradizionali fasi di fetch ed execute, allargando la visibilità sulle istruzioni da eseguire e applicando tecniche di scheduling ottimizzato. Ciò richiede principalmente che la fase di decodifica istruzioni del Pentium Pro sia più intelligente in termini di predizione del flusso di programma. Lo scheduling ottimizzato richiede che la fondamentale fase di "execute" sia sostituita da fasi disgiunte di "dispatch/execute" e di "retire". Ciò permette alle istruzioni di essere iniziate in qualsiasi ordine ma di essere sempre completate nella sequenza originale del programma in esecuzione. Il Pentium Pro utilizza tre indipendenti motori accoppiati con un'Instruction Pool, come mostrato nello schema in questa pagina.

Il problema principale

Prima di scendere maggiormente nei dettagli su come il Pentium Pro raggiunge le sue elevate performance, consideriamo il segmento di codice mostrato qui di seguito:

- (1) R1 := mem[R0]
- (2) R2 := R1 + R2
- (3) R5 := R5 + 1
- (4) R6 := R6 - R3

La prima istruzione carica nel registro

R1 la cella di memoria indirizzata da R0, la seconda somma al contenuto di R2 il contenuto di R1, la terza incrementa R5 e la quarta sottrae il valore di R3 al contenuto di R6.

Ipotizziamo che a tempo di esecuzione, la prima istruzione provochi un cache miss (mancanza del dato richiesto all'interno della cache). Una CPU di concezione tradizionale – dal semplice 8088 all'evoluto Pentium "liscio" – non può far altro che attendere pazientemente che l'interfaccia di bus provveda a richiedere alla memoria principale il dato mancante, provocando un rallentamento del sistema.

Se consideriamo che la velocità delle CPU è aumentata dieci volte negli ultimi dieci anni, mentre la velocità delle memorie è migliorata solo del 60 per cento è facile rendersi conto che un microprocessore costruito secondo schemi tradizionali passerebbe buona parte del suo tempo ad attendere risposte dalla memoria. Tale "rallentamento" delle memorie (rispetto alla velocità sempre crescente delle CPU) è il problema che il Pentium Pro cerca di risolvere.

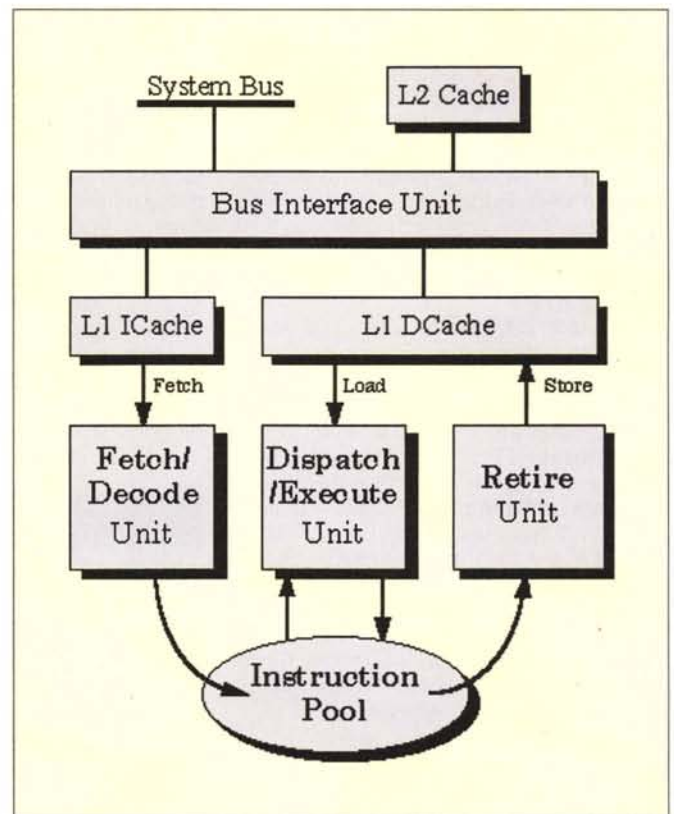
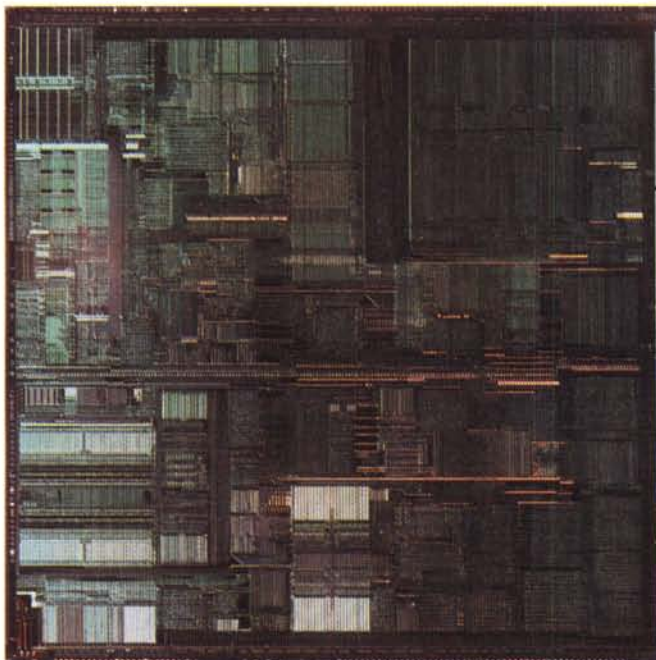
Un primo approccio risolutivo potrebbe consistere nell'utilizzare uno specifico chipset separato che si occupi dell'ottimizzazione del codice a tempo di esecuzione. Ma una CPU, ad alte pre-

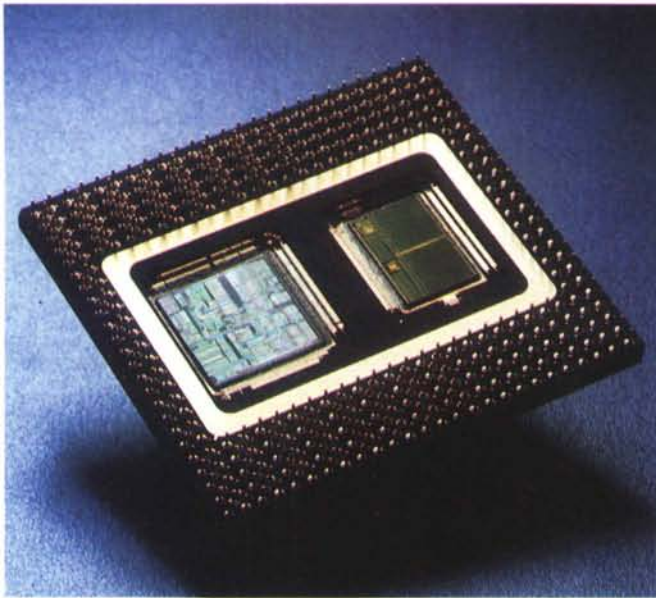
stazioni, che necessiti di componenti specializzati di supporto ad altissima velocità non rappresenterebbe una soluzione economicamente valida per sistemi prodotti in grandi volumi. Un secondo approccio al problema, certamente più brutale, può consistere nell'aumentare a dismisura le dimensioni della cache di secondo livello in modo da minimizzare i cache miss. Sebbene funzionante, questa è un'altra soluzione costosa, specialmente considerando le esigenze di velocità dei componenti per le cache di secondo livello. La soluzione proposta per il Pentium Pro sfrutta l'alta tecnologia concentrata tutta all'interno del chip, in modo da utilizzare memorie ed architetture esterne di tipo tradizionale e quindi anche piuttosto economiche.

Tecnica look ahead

Per evitare il problema dei ritardi di memoria, il Pentium Pro "guarda avanti" nel suo Instruction pool alle istruzioni successive lavorando sempre attivamente piuttosto che attendere passivamente il verificarsi degli eventi. Nell'esempio del codice prima mostrato la seconda istruzione non è eseguibile finché non si ha il risultato della prima istruzione (il valore finale di R2 dipende

A sinistra una fotografia del "silicio" del Pentium Pro, a destra lo schema a blocchi.





Il Pentium Pro è il primo processore Intel a incorporare nello stesso contenitore ceramico sia la CPU vera e propria che la cache di secondo livello. La sola sezione microprocessore incorpora circa 5.5 milioni di transistor.

di clock e sono supportati quattro accessi concorrenti (uno in ogni stadio della sua pipeline). Se il nucleo della CPU effettua un quinto accesso di lettura alla cache, questo viene bufferizzato in attesa della sua esecuzione. Anche in questo caso, però, la CPU non si arresta per attendere l'esito della sua richiesta. Infatti il load buffer della cache può mantenere fino a dodici richieste prima di provocare un arresto del microprocessore.

Gli accessi che non ottengono esito all'interno della cache L2 (cache miss) sono passati al sottosistema interfacciato con la memoria centrale. L'interfaccia bus del Pentium Pro può effettuare fino a quattro richieste alla memoria per sistemi monoprocesso e fino ad otto per sistemi a più processori: tutte le transazioni con la memoria avvengono utilizzando il protocollo MESI per mantenere coerente il contenuto delle varie cache nei sistemi multiprocessore. Da alcuni benchmark simulati effettuati da Intel durante la progettazione del Pentium Pro è stato misurato che il bus esterno del processore (grazie, prevalentemente, alla presenza della cache di secondo livello integrata) è utilizzato per meno del 25% della sua larghezza di banda. Ciò significa che fino a quattro Pentium Pro possono essere facilmente utilizzati insieme per realizzare sistemi multiprocessor senza avere troppi problemi sul bus. A tutto vantaggio della progettazione di sistemi multiprocessor che presto potranno finire sulla nostra scrivania, per offrire performance sempre più elevate.

Dalla teoria alla pratica

Tutto quello che avete letto finora è, ovviamente, basato sulla documentazione Intel del nuovo processore e naturalmente non abbiamo motivi di dubitare sulla sua veridicità. Altrettanto ovviamente, però, non potevamo evitare di effettuare qualche test comparativo con altri processori per capire quanto l'elevata

anche da R1), ma lo sono la terza e la quarta. Il Pentium Pro (in maniera "speculativa") esegue queste due istruzioni durante l'attesa del dato richiesto in memoria.

Il cache miss dell'istruzione 1 può provocare un'attesa anche di diversi cicli di clock durante i quali il Pentium Pro continua a cercare altre istruzioni da eseguire, cercando anche 20 o 30 istruzioni più avanti all'interno del codice. La tecnologia dell'esecuzione dinamica in pratica effettua un'ottimale aggiustamento dell'esecuzione delle istruzioni predicendo il flusso del programma, analizzando il grafo del flusso di dati per scegliere il miglior ordinamento di esecuzione delle istruzioni. Il Pentium Pro dinamicamente ottimizza il suo lavoro, definito dal flusso di istruzioni in entrata, per minimizzare il tempo di esecuzione globale.

Le unità interne

L'unità "fetch/decode" del Pentium Pro è incaricata di inviare le istruzioni all'Instruction Pool. Per fare questo preleva le istruzioni a partire da quella "puntata" dal program counter e le analizza per la loro decodifica. Se all'interno del codice è incontrato un salto condizionato la sua destinazione (l'esito della condizione) è calcolata in anticipo e il prelievo successivo delle istruzioni continua da questo punto. Il Pentium Pro è in grado di predire l'esito di questo salti correttamente con probabilità di successo superiore al 90 per cento.

L'unità "dispatch/execute" è incaricata di "osservare" il codice da eseguire. Quando questa trova un'istruzione

che ha tutti i suoi operandi pronti la invia all'unità esecutiva. La logica di funzionamento, di tipo data-flow, offre una riduzione del tempo di esecuzione complessivo di un programma scegliendo di volta in volta tutte le istruzioni "pronte" per l'esecuzione. In questo modo le istruzioni "lente" non bloccano le istruzioni "veloci".

L'unità "Retire" individua nell'Instruction Pool quelle istruzioni che sono state completate e passa i risultati allo stato globale della macchina ricostruendo l'ordine del programma originale.

L'unità di fetch/decode lavora secondo l'ordine del programma e può inviare fino a tre istruzioni all'Instruction Pool per ogni ciclo di clock. L'unità dispatch/execute opera "fuori ordine" e può processare fino a cinque istruzioni per ciclo di clock con una media di tre. Lo stesso succede (tre istruzioni per clock) per l'unità "Retire".

Le cache (dati e istruzioni) sono di tipo dual port e sono in grado di eseguire una lettura e una scrittura in ogni ciclo di clock. Il throughput è di quattro cicli

Pentium Pro: versioni e caratteristiche

Velocità (MHz)	150	166	180	200
Velocità Bus (MHz)	60	66	60	66
L2 Cache	256K	512K	256K	256K
Processo	0.6 micron	0.35 micron	0.35 micron	0.35 micron
Dimensioni	691 mils/side	552 mils/side	552 mils/side	552 mils/side
Voltaggio	3.1V	3.3V	3.3V	3.3V
Corrente massima	9.9	9.4	10.1	11.2
Potenza (Watt)	23.2 typ; 29.2 max	N/A	25.3 typ; 31.7 max	28.1 typ; 35.0 max
Performance	6.08 SPECint95 5.42 SPECfp95	7.11 SPECint95 6.21 SPECfp95	7.29 SPECint95 6.10 SPECfp95	8.09 SPECint95 6.70 SPECfp95

tecnologia utilizzata nel Pentium Pro sia effettivamente fruibile da parte dell'utente in termini di maggiori prestazioni.

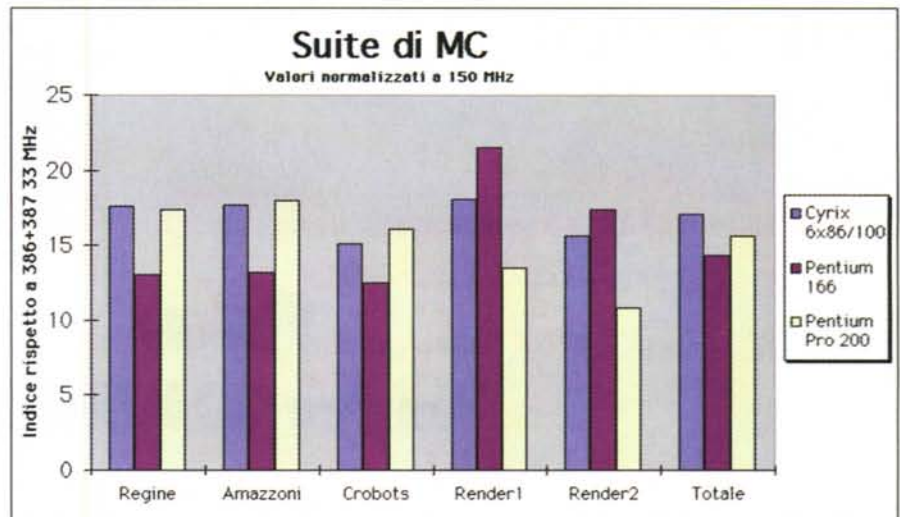
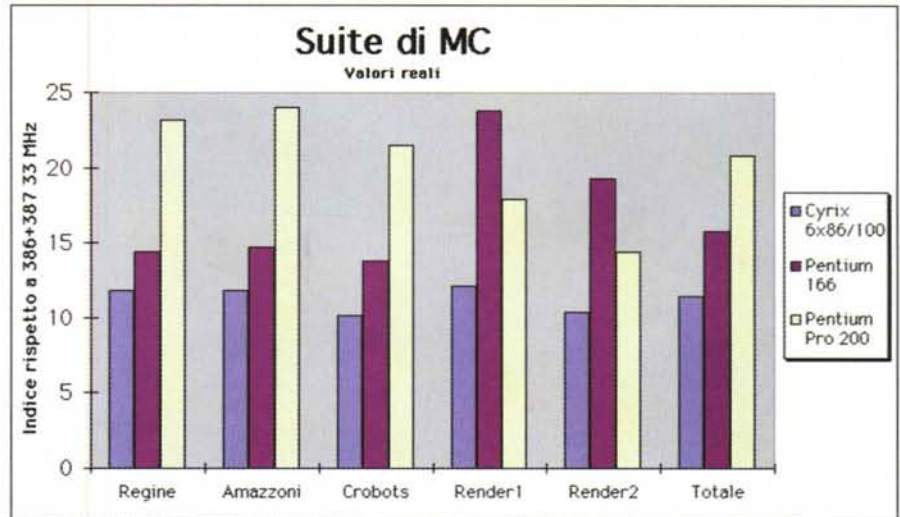
Il benchmark utilizzato è, ancora una volta, l'ormai nota Suite di MC: una collezione di programmi particolarmente "tosti" dal punto di vista elaborativo, ognuno indirizzato al testing di particolari proprietà dei microprocessori: aritmetica intera, floating point, interfacciamento con cache e con la memoria centrale.

I chip di paragone scelti sono stati il Pentium "normale" e... quella spina nel fianco di Mamma Intel di nome 6x86 e cognome Cyrix: un interessante microprocessore che, nonostante la sua architettura "quasi tradizionale" (si tratta, potremmo dire, di un Pentium particolarmente ottimizzato), nasce col preciso intento di far concorrenza ai chip di sesta generazione.

Nel primo grafico mostrato in questa pagina sono riportati i risultati reali ottenuti dai vari microprocessori. Il Pentium Pro ha il clock a 200 MHz, il Pentium a 166 e il Cyrix a 100. Colpo di scena: il Pentium Pro a 200 MHz è, in alcuni casi, più lento del suo fratello minore a 166! È successo nei test Render1 e Render2 che utilizzano in maniera massiccia il coprocessore matematico integrato all'interno del chip.

Non gridate allo scandalo: è d'obbligo una precisazione. La nostra Suite è compilata a 16 bit (al fine di rispecchiare al meglio l'attuale disponibilità di software prevalentemente di questo tipo) e non è affatto detto che i risultati sarebbero dello stesso tipo con applicazioni a 32. Ciò dimostra, semplicemente, che il Pentium è più ottimizzato per il codice "vecchio" mentre per sfruttare al massimo le elevate performance del Pentium Pro è necessario far girare codice "nuovo" (a 32 bit).

E il Cyrix che fa, resta indietro a guardare? Non proprio. Non dimentichiamo, infatti, che il confronto è piuttosto impari: il Cyrix (nella versione da noi testata) è a 100 MHz (ma già esiste anche il 133 MHz), mentre il Pentium Pro è a 200. Bene, tiriamo fuori la bacchetta magica e normalizziamo i valori raggiunti dai tre differenti processori ad una velocità media di 150 MHz. Non si tratta di un'operazione scientificamente lecita (ne siamo ben consci...) ma può essere utile per avere comunque qualche informazione in più. Seconda sorpresa: Cyrix batte tutti nel risultato totale dell'intera Suite, eguaglia sostanzialmente il Pentium Pro per quel che riguarda i calcoli in aritmetica intera (primi tre test della Suite), lo batte in larga misura per quel che riguarda i calcoli in floating point pur non rag-



giungendo (in questo campo) gli strabilianti risultati del Pentium "liscio". Incredibile, ma vero!

Tirando le somme

Una prima, interessante, conclusione riguarda il fatto che per sfruttare appieno le potenzialità offerte dal Pentium Pro è assolutamente indispensabile utilizzare software a 32 bit. Naturalmente la considerazione non riguarda solo i programmi applicativi, ma anche (o soprattutto) il sistema operativo funzionante sulla macchina. Scordatevi, se non volete buttare tempo e denaro, la possibilità di installare su una macchina Pentium Pro il vecchio Windows 3.1 o addirittura l'arcaico MS-DOS. Rischiereste di avere performance globali del tutto paragonabili a quelle offerte da macchine più economiche senza alcun vantaggio effettivo dato

dalla maggiore sofisticazione tecnologica. Si parte, come minimo, da Windows 95, passando poi per OS/2, Solaris, Linux e altri sistemi operativi full 32 bit. Se provenite dalla vecchia scuola (e li intendete restare), lasciate perdere il Pentium Pro. Meglio, sicuramente, optare per altre soluzioni hardware meno avverse al trattamento del codice "mezzo formato".

Un secondo spunto di riflessione, riguarda la tecnologia di costruzione dei moderni microprocessori che consente frequenze di clock più o meno elevate. Intel è già a quota 200 MHz, ma non è detto che Cyrix non sia in procinto di proporre il suo microprocessore alla stessa velocità. Abbiamo visto dai benchmark quanto sia ottimizzato il 6x86, penalizzato attualmente solo dalla (relativamente) bassa frequenza di clock. La partita è aperta: Signori Costruttori, fate il vostro gioco... MS