



# LA LEGGE DI MOORE E LO SVILUPPO DEI CIRCUITI INTEGRATI

Si analizzano i fattori tecnici ed economici che hanno consentito lo sviluppo esponenziale della microelettronica su un arco temporale di oltre 30 anni e si descrivono le iniziative in atto al fine di consentire il mantenimento di questo sviluppo nei prossimi 10-15 anni. Si discutono poi nuovi scenari che potrebbero consentire aumenti di prestazioni paragonabili a quelli che la microelettronica ha già fornito e sta ancora fornendo.

Livio Baldi  
Gianfranco Cerofolini

## 1. LA LEGGE DI MOORE

In scienza, con la definizione “legge” si intende “un principio fondamentale di fenomeni naturali [...] tratto dalla constatazione del costante verificarsi di un effetto in dipendenza di determinate cause” [1].

L’osservazione di Moore del 1975, sul fatto che il numero di *transistor* per circuito integrato era aumentato dalla data della loro invenzione, di un fattore 2 ogni anno su un arco temporale di 16 anni [2], ben difficilmente, può essere considerata una legge nel senso sopra citato. Malgrado ciò, la persistenza di tale crescita esponenziale negli anni successivi (seppure con un “tempo di raddoppiamento” che si è andato assestando sui 18 mesi) è stata così regolare (Figura 1) che l’attribuzione “legge di Moore” non è stata posta in discussione.

L’iniziale analisi di Moore aveva indicato come fattori (in ordine di importanza crescente) contribuenti allo sviluppo esponenziale della complessità:

■ l’aumento di area del *chip* (la frazione di cristallo di silicio su cui viene realizzato il circuito integrato);

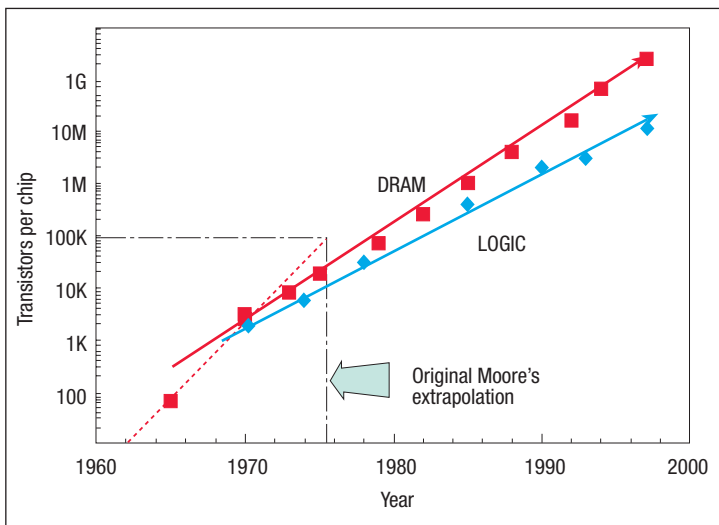
■ la riduzione delle dimensioni delle strutture elementari che componevano i circuiti integrati;

■ il miglioramento della capacità di progettare dispositivi (siano essi **transistor bipolari o di tipo MOS**, *Metal Oxide Semiconductor Field Effect Transistor*) e circuiti più efficienti. Moore stesso riconosceva che alcuni di questi fattori avrebbero avuto un impatto limitato e, quindi, l’evoluzione successiva sarebbe dipesa essenzialmente dal solo ridursi delle dimensioni critiche e, di conseguenza, avrebbe rallentato il suo ritmo.

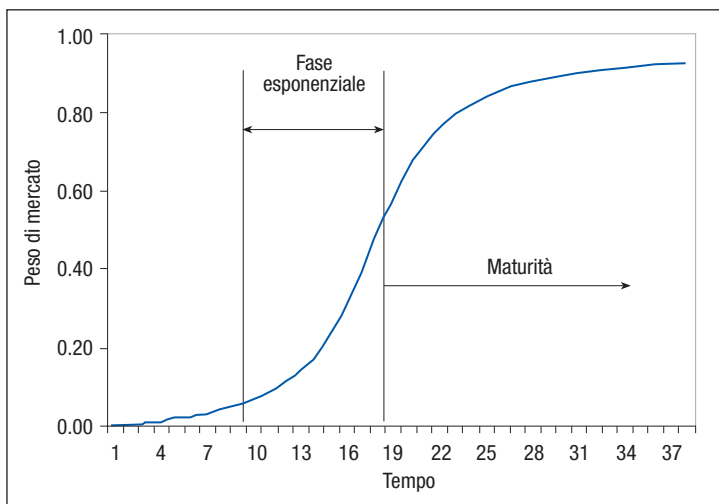
L’analisi è stata sostanzialmente confermata dagli eventi degli anni successivi:

■ l’area del *chip* è aumentata di un fattore 4 dal 1965 al 1975, grazie ad un generale miglioramento della qualità dei materiali, ma è poi stata limitata dai costi accettabili dei prodotti (il costo per unità di area dei circuiti integrati, espresso in dollari/cm<sup>2</sup>, è rimasto sostanzialmente stabile nel tempo);

■ la dimensione lineare minima che viene definita sulle strutture MOS si è ridotta di un fattore 3 ogni 10 anni, ed ha fornito il contributo principale;



**FIGURA 1**  
Prima legge di Moore sull'evoluzione della densità (in transistor/dispositivo) dei circuiti integrati



**FIGURA 2** | il miglioramento nella capacità di progettare dispositivi complessi si è concentrato essenzialmente sul tempo di progettazione, con l'introduzione di tecniche automatizzate (librerie di funzioni logiche, sintesi logica), a scapito della densità dei circuiti;

per le memorie, tecniche di auto-allineamento e di sviluppo tridimensionale dei dispositivi, hanno permesso di ridurre ulteriormente l'ingombro.

## 2. I FATTORI CHE HANNO SOSTENUTO LA LEGGE DI MOORE

Comprendere le cause di uno sviluppo esponenziale non è difficile. Ogni nuovo prodotto si espande occupando nicchie via via più am-

pie. Qualunque famiglia di prodotti si consideri, la curva che ne descrive la progressiva crescita è di regola una curva a "S" (Figura 2) caratterizzata da una fase iniziale esponenziale e da una successiva tendenza alla saturazione (corretta verso l'alto dalle modifiche tecnologiche marginali apportate al prodotto maturo). Le fasi iniziali di questo tipo di crescita sono esponenziali in quanto l'occupazione della nicchia dà un vantaggio proporzionale al grado di occupazione. Se l'utile che ne deriva viene reinvestito nella produzione di un nuovo bene, e se l'efficacia di questa operazione è proporzionale al reinvestimento, allora l'equazione che regola il fenomeno, integrata, porta ad una crescita esponenziale.

Per quanto riguarda la microelettronica, il fattore sorprendente è che il parametro che si sta considerando non è un fattore economico, come l'occupazione di mercato, ma un fattore prettamente tecnologico, come la densità di integrazione. Per spiegare il fenomeno, bisogna quindi, introdurre alcune ipotesi aggiuntive:

1. il costo di produzione è proporzionale all'area dei circuiti integrati;
2. il mercato è in grado di recepire una quantità pressoché illimitata di prodotti;
3. l'aumento della densità di integrazione è determinato solo dalle risorse investite in ricerca e sviluppo, cioè non esistono barriere tecnologiche fondamentali.

Il primo fattore è determinato dal fatto che i circuiti integrati sono prodotti in serie, su substrati di silicio di dimensioni standard: il wafer (Figura 3). Ogni wafer ospita un numero di dispositivi dettato dal rapporto tra area del wafer e area del circuito integrato. Dato che il costo di lavorazione di ogni wafer di silicio è generalmente costante (in realtà, cresce lentamente nel tempo con l'aumento della complessità della tecnologia), il costo del circuito integrato è inversamente proporzionale all'area.

Per quanto riguarda il secondo fattore, bisogna considerare che i circuiti integrati hanno consentito di sostituire a minor costo circuiti formati da tanti dispositivi discreti assemblati su di una piastra stampata. E poiché in questa logica è possibile sostituire circuiti prima semplici, poi complessi, quindi, appa-

rati, sottosistemi e, infine, sistemi, i circuiti integrati hanno finito con l'occupare praticamente tutto il mercato dell'elettronica.

In effetti poi, piuttosto che agire su un mercato di sostituzione (dalla radio o televisore a valvole ai corrispondenti apparati "a transistor", e poi a circuiti integrati), la drammatica riduzione di costi ed aumento di complessità derivanti dal processo di integrazione crescente hanno consentito la realizzazione di nuovi prodotti o, comunque, la produzione di volume (a costi tali da rendere il bene accessibile ad un vastissimo pubblico) di prodotti già esistenti, ma limitati a pochissimi utenti. A testimonianza di ciò, e senza nessuna presunzione di completezza, basta citare la radio portatile, la televisione a colori, il telecomando, il fax, il cellulare, le fotocamere e videocamere digitali, il navigatore satellitare ecc.. La crescita esponenziale del mercato della microelettronica e dell'elettronica, è quindi, in realtà, dovuto al sovrapporsi di tante curve a "S", relative a varie famiglie di prodotti.

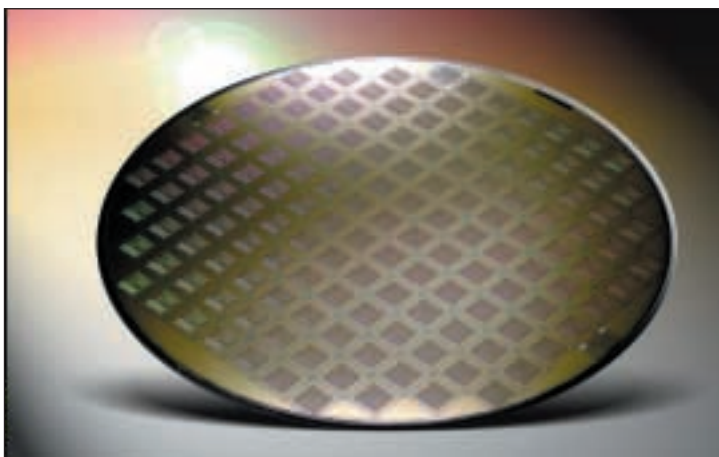
Il terzo fattore è quello più sorprendente, poiché la crescita esponenziale si è mantenuta invariata per quasi 40 anni dando luogo ad un aumento di produttività senza precedenti nella storia dell'umanità, e richiede una spiegazione a parte.

### 3. L'ELEMENTO BASE DEI CIRCUITI INTEGRATI E LA SUA SCALABILITÀ

Per capire su quali basi tecniche sia (stata) possibile la "rivoluzione tranquilla" della microelettronica, bisogna considerare l'elemento base che costituisce i circuiti integrati: il **transistor MOS** (si veda il box di pag. 15).

Gran parte del continuo successo della tecnologia CMOS nel tempo è legato ai seguenti fattori:

1. la semplicità intrinseca della struttura (almeno agli inizi);
2. la possibilità, grazie all'architettura CMOS, di realizzare circuiti integrati che dissipano solo durante il funzionamento, e non in quiescenza, con evidenti risparmi di energia;
3. la scalabilità intrinseca dei transistor, che ha reso possibile la crescita di complessità e prestazioni dei sistemi, senza richiedere innovazioni significative della fisica di base;
4. la disponibilità di soluzioni tecnologiche



che hanno consentito di realizzare strutture via via più piccole.

Espressa in termini elementari, la scalabilità consiste nel fatto che, se un transistor di certe dimensioni funziona, è possibile costruirne uno più piccolo e funzionante nella stessa maniera (almeno dal punto di vista del campo elettrico) abbassando dello stesso fattore i potenziali esterni ed aumentando del fattore inverso i drogaggi (**scaling**, vedi pag. 6) [3]. Ovviamente questa proprietà vale sino a che non intervengano limiti fisici fondamentali.

Tra i fattori tecnologici, il principale è stato la possibilità di definire su silicio strutture sempre più piccole. La tecnica base su cui è fondata la definizione delle strutture elementari, la **fotolitografia** (vedi pag. 7) non ha incontrato problemi nel ridurre le dimensioni minime definite finché queste non sono diventate confrontabili con la lunghezza d'onda della radiazione (ha consentito, quindi, la riduzione da  $50 \mu\text{m}$  a  $0.5 \mu\text{m}$ ). Nonostante la fine della litografia "ottica", basata sull'uso di lenti o specchi per proiettare sul silicio l'immagine della struttura da realizzare, sia stata prevista più volte, e con ottime motivazioni teoriche, le tecnologie alternative (come la litografia a raggi X, o l'uso di fasci di elettroni o di ioni per definire direttamente le strutture) non hanno mai raggiunto maturità industriale. È risultato, invece, vincente l'approccio evolutivo, basato sulla riduzione della lunghezza d'onda della radiazione utilizzata (adesso giunta a  $157 \text{ nm}$ ), sullo sfruttamento dei fenomeni di diffrazione e di interferenza della luce (*phase shift masks* e *optical proximity correction*) e sullo sviluppo di sistemi ottici complessi (illuminazione fuori asse) e ad alta apertura numerica, che hanno consentito di definire

**FIGURA 3**

*Un substrato di silicio (wafer), qui rappresentato al termine della lavorazione, costituisce l'unità base della produzione in microelettronica*

Le grandi dimensioni del transistor, rispetto a quelle per cui gli effetti quantistici divengono importanti, hanno consentito di descrivere il transistor attraverso un insieme di 5 equazioni differenziali (l'equazione di Poisson, le equazioni di trasporto e di continuità per elettroni e lacune). Dall'equazione di Poisson si deduce un teorema di scala che è stato cardine nello sviluppo della microelettronica [4]. Questo teorema di scala stabilisce che la distribuzione del campo elettrico all'interno di un transistor non cambia se scalando le dimensioni lineari di un fattore  $K$  in una dimensione (per esempio la lunghezza di gate  $L$ ) si scalano tutte le altre dimensioni (per esempio: lo spessore dell'ossido di gate  $t_{ox}$ , la larghezza del transistor  $w$  e la profondità di giunzione  $x_j$ ) dello stesso fattore e contemporaneamente si cambiano le condizioni al contorno dello stesso fattore e la densità di drogante del fattore inverso  $1/K$  (Tabella).

Il controllo dei parametri fisici è stato reso possibile dalla continua evoluzione tecnologica, senza particolari traumi, mentre la riduzione dei potenziali esterni è stata a lungo evitata, sia per problemi di standardizzazione delle tensioni di alimentazioni, sia perché, come evidenziato in tabella, scalando a *tensioni* costante, anziché a *campo elettrico* costante, si ottengono notevoli vantaggi nelle prestazioni di velocità. Tuttavia, nel caso di scaling a tensione costante, si verifica un preoccupante incremento della densità di corrente e di potenza, e dei campi elettrici, il che tende a compromettere l'affidabilità dei dispositivi, facilitando fenomeni di elettromigrazione nelle interconnessioni e di degrado degli ossidi di gate.

Per questi motivi, a partire da lunghezze di gate da  $0,5 \mu\text{m}$ , si è iniziato a scalare la tensione di alimentazione, che prima era stata mantenuta costante a 5 V, in modo proporzionale alla lunghezza di canale. Questo tuttavia ha comportato numerose difficoltà, in quanto diverse componenti dei circuiti integrati complessi, per esempio le memorie *flash*, quelle dinamiche ad accesso casuale e i circuiti analogici, richiedono tensioni minime di alimentazione superiori a quelle dettate dalle **regole di scaling**. Per questo motivo le tecnologie CMOS avanzate comprendono in genere diversi tipologie di transistor, ottimizzati per diverse tensioni di alimentazione.

Parametri di progetto	Approccio generalizzato	Campo costante	Tensione costante
Dimensioni ( $w, L, t_{ox}$ )	$1/\lambda$	$1/\lambda$	$1/\lambda$
Drogaggio	$\lambda^2/K$	$\lambda$	$\lambda^2$
Tensioni	$1/K$	$1/\lambda$	1
<b>Variabili dipendenti</b>			
Campo Elettrico	$\lambda/K$	1	$\lambda$
Correnti	$\lambda/K^2$	$1/\lambda$	$\lambda$
Capacità	$1/\lambda$	$1/\lambda$	$1/\lambda$
Dissipazione di Potenza	$\lambda/K^3$	$1/\lambda^2$	$\lambda$
Ritardo	$K/\lambda^2$	$1/\lambda$	$1/\lambda^2$
Prodotto Ritardo $\times$ Consumo	$1/\lambda K^2$	$1/\lambda^3$	$1/\lambda$
Densità di potenza	$\lambda^3/K^3$	1	$\lambda^3$

strutture più piccole della lunghezza d'onda. È opinione corrente che, con le tecnologie ora in fase di industrializzazione, sarà possibile definire transistor da 65 nm, mentre, per andare oltre, è in avanzato sviluppo l'utilizzo di lunghezze d'onda da 13 nm (ultravioletto estremo o raggi X molli), che dovrebbero assicurare ancora notevoli margini di crescita.

La crescita è stata facilitata dal fatto che in realtà, i primi dispositivi MOS avevano dimensioni di molto superiori ai limiti fisici (lunghezza d'onda della radiazione usata per la litografia, lunghezze di diffusione dei portatori di carica nel substrato, dimensioni per le quali gli effetti quantistici diventano dominanti, passo del reticolo del silicio), e determinate essenzialmente dalle capacità tecnologiche e dalla qualità dei materiali disponibili.

Con la riduzione progressiva delle dimensioni sono emersi vari ostacoli tecnologici, lega-

ti al peso crescente di effetti parassiti, che sono stati superati grazie ad una serie di innovazioni nel campo dei materiali e dei processi produttivi. Si possono citare di seguito:

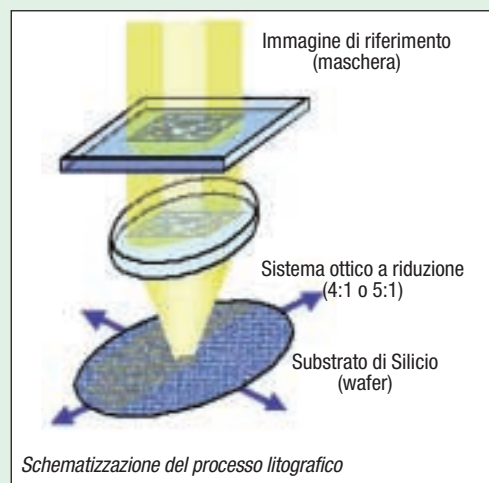
- il miglioramento della qualità dei substrati di silicio, le cui dimensioni sono passate dal valore iniziale di 25.4 mm (1 inch) agli attuali 300 mm, con notevoli riduzioni nei costi di produzione;

- l'introduzione di tecnologie autoallineate (note con gli acronimi LOCOS, *Local-Oxidation of Silicon*, o PLANOX, rispettivamente messe a punto in Philips o SGS) basate su di uno strato protettivo di nitrato di silicio, per isolare elettricamente tra di loro i transistor;

- l'uso del silicio policristallino depositato da fase gassosa per realizzare il gate dei transistor (tecnologia *silicon gate*, messa a punto in INTEL) che ha permesso di ottenere MOS con *source* e *drain* autoallineato;

La litografia è la tecnologia fondamentale per la microelettronica, in quanto permette di definire su silicio le strutture elementari che formeranno i transistor e le loro interconnessioni.

Il principio è abbastanza semplice, abbastanza simile ad una stampa fotografica a riduzione, ma ciò che rende il procedimento difficile sono le ridottissime dimensioni in gioco (frazioni di *micron*). Nelle stampe fotografiche, un'immagine di riferimento (il negativo) viene proiettata da un sistema ottico (in genere, ingrandendola) su di un substrato fotosensibile. Sviluppando quest'ultimo si ottiene una copia dell'immagine. Nel caso della **fotolitografia**, l'immagine di riferimento è costituita da una "maschera", cioè una lastra di quarzo su cui sono state disegnate in materiale opaco, le strutture che si vogliono trasferire su silicio. Le strutture su maschera hanno in genere dimensioni 4 o 5 volte superiori al valore finale, e sono realizzate con un fascio di elettroni. Un sistema ottico complesso (l'insieme delle lenti può arrivare a pesare qualche tonnellata!) proietta l'immagine, riducendone le dimensioni, sul substrato di silicio da lavorare (wafer) (Figura). Questo è stato preventivamente ricoperto da una resina, un materiale in genere organico, fotosensibile; la parte della resina che è stata illuminata cambia composizione, e può essere facilmente rimossa senza intaccare la parte non illuminata. La resina rimasta, indurita da opportuni trattamenti, viene poi utilizzata per proteggere gli strati sottostanti da attacchi chimici o fisici, che rimuovono gli strati superficiali, o per schermare il silicio sottostante da materiali droganti, introdotti con una forma di bombardamento ionico. La scelta di ridurre l'immagine è stata fatta sia per semplificare la fabbricazione della maschera, sia per ridurre il peso dei difetti (granelli di polvere, imperfezioni nella fabbricazione della maschera ecc.). Per utilizzare al meglio il sistema di lenti, e per contenere le dimensioni delle maschere, queste comprendono solo pochi dispositivi, e coprono solo una porzione ridotta del wafer di silicio, dell'ordine di qualche centimetro quadrato. La copertura di tutto il wafer viene realizzata mediante esposizioni multiple, spostando il wafer rispetto al sistema di esposizione in modo controllato. Per dare un'idea delle dimensioni in gioco, il supporto meccanico mobile dei wafer, ha una precisione di posizionamento dell'ordine del milionesimo di millimetro, e si muove con una accelerazione di 6 gravità.



- l'introduzione della impiantazione ionica per drogare, in modo controllato, attraverso l'energia del fascio e la dose totale impiantata, il silicio su cui sono realizzati i transistori, il che ha permesso di ridurre le tensioni di alimentazione sotto i 5 V;

- l'introduzione di architetture speciali (*drain extension*) per ridurre i campi elettrici massimi nei transistor, che hanno permesso di realizzare transistor con lunghezza di canale inferiore al micron;

- l'introduzione da parte di IBM, in tempi recenti, della lappatura chimico-fisica per la planarizzazione delle superfici, che ha consentito la realizzazione di molteplici livelli di interconnessione (sino a 7-8) e di utilizzare il rame come conduttore su chip, riducendo i problemi di affidabilità.

#### 4. QUANTO A LUNGO ANCORA?

La continuazione dello sviluppo esponenziale della tecnologia microelettronica è un elemento fondamentale per la crescita esponenziale del mercato della microelettronica, e della elettronica più in generale, in quanto è proprio la disponibilità di transistor, e quindi di memoria e di capacità di elaborazione

dell'informazione, a prezzi sempre più bassi, che rende possibile introdurre continuamente nuove applicazioni e aprire nuovi mercati.

La domanda se esistano limiti prevedibili a questo sviluppo riveste quindi un interesse fondamentale, non solo per la tecnologia, ma anche, più in generale per l'economia.

Possiamo distinguere tra tre tipi di limite: fisici, tecnologici e economici.

##### 4.1. I limiti fisici

Quali fattori limiteranno prima o poi lo sviluppo della microelettronica dal punto di vista tecnologico è, tuttora, oggetto di intensa discussione. Due possibili fattori tecnici sono: la non scalabilità del transistor MOS sotto una certa dimensione  $L_{phys}$  (il "limite fisico") e l'impossibilità tecnologica di definire geometrie sotto una dimensione critica  $L_{litho}$  (il "limite litografico"). Tra i vari lavori che discutono il limite fisico si citano quelli di Packan [4], Keyes [5] e Frank et al. [6]; tra quelli che discutono i limiti della litografia si citano i lavori di Ito e Okazaki [7] e di Harriott [8].

Una lucida analisi, prodotta recentemente da Lloyd [9], sembra indicare che i limiti di natura fondamentale siano lontanissimi, lasciando



con ciò aperta la prospettiva dello sviluppo di nuove tecnologie. Concludendo, la sua analisi, Lloyd afferma: “il computer di cui sto scrivendo fornisce forti evidenze sperimentali che, se la teoria fisica di base è corretta, supportano la possibilità di un computer universale. Che sia o no possibile effettuare la computazione nei regimi estremi dipende dagli sviluppi tecnologici futuri, difficili da predire. Se, come peraltro sembra altamente improbabile, è possibile mantenere in futuro il progresso esponenziale della legge di Moore, ci vorranno solo 250 anni per coprire i 40 ordini di grandezza che separano le prestazioni dei computer attuali ( $10^{10}$  operazioni al secondo su  $10^{10}$  bit) dalle prestazioni dell'ultimo laptop computer (che, del peso di 1 kg, effettua  $10^{51}$  operazioni al secondo su  $10^{31}$  bit)” [9].

Va da sé che allo stato attuale, non ha senso ipotizzare percorsi tecnologici con cui avvicinarsi a tali prestazioni. Peraltro, su un arco temporale più ridotto, e con obiettivi certamente più ridimensionati, si possono individuare iniziative o architetture che possono contribuire in maniera significativa ad aumentare il livello di complessità dei circuiti integrati. Infatti transistor MOS con dimensioni di 25 nm sono già stati dimostrati sperimentalmente [10], assicurandoci ancora almeno 5 generazioni di sviluppo tecnologico.

#### 4.2. I limiti tecnologici

Più difficile è garantire la realizzabilità pratica di queste tecnologie avanzate, che dipendono dalla disponibilità di un'ampia rete di supporto, costituita da attrezzature per la fabbricazione, materiali di estrema purezza, sistemi di misura e dal software per la progettazione di circuiti integrati complessi.

Per evitare che l'accrescimento esponenziale descritto dalla legge di Moore rallenti (rendendo quindi matura l'industria microelettronica e riducendo il positivo impatto di questa sull'intero comparto produttivo) si è avviata un'attività volta a:

- individuare i colli di bottiglia che limitano lo sviluppo della microelettronica secondo la legge di Moore;

- suggerire le attività di ricerca necessarie a rimuovere tali colli di bottiglia.

Si è passati quindi da una fase di “evoluzione esponenziale spontanea” ad una di evoluzio-

La **International Technology Roadmap for Semiconductors** (<http://public.itrs.net>) deriva dallo sforzo collaborativo di diverse centinaia di ricercatori industriali (operanti sia nella produzione di circuiti integrati che di attrezzature) e accademici, sparsi in tutto il mondo. La sua origine è da ricondursi alla NTRS (*National Technology Roadmap for Semiconductors*) roadmap, rilasciata per la prima volta dalla SIA (*Semiconductor Industry Association*) statunitense nel 1992, con lo scopo di “mantenere il tasso di sviluppo storico su di un orizzonte temporale di 15 anni”. In seguito, nel 1998, vennero invitati a partecipare anche industrie e centri di ricerca europei, coreani, giapponesi e taiwanesi. Nel 1998, venne prodotto congiuntamente un aggiornamento della NTRS roadmap del 1997, e in seguito si è proceduto a rilasciare nuove versioni negli anni dispari, con aggiornamenti negli anni pari. Le varie aree tecnologiche vengono contrassegnate in bianco, giallo o rosso a seconda che esistano già soluzioni disponibili, che ci siano soluzioni possibili, richiedenti solo un certo lavoro di sviluppo, oppure che non sia nota nessuna soluzione. Per questo motivo le tecnologie più avanzate sono designate come “il muro di mattoni rossi”, perché predomina l'incertezza. L'ultima edizione della roadmap, datata 2001, si spinge fino al 2014, data per la quale è prevista la tecnologia MOS da 30 nm [5].

ne guidata, rappresentata dalla ITRS: **International Technology Roadmap for Semiconductors** [11].

La roadmap altro non è che una descrizione dettagliata dello scenario necessario a supportare la crescita esponenziale della densità, diviso nei suoi componenti fondamentali (dimensioni dei dispositivi, caratteristiche elettriche fondamentali, tensione di alimentazione, correnti dissipate, caratteristiche dei materiali), con l'indicazione delle tecnologie da sviluppare e delle date per cui devono essere disponibili.

Questa roadmap è divenuta ormai anche uno strumento di lavoro per organizzare le attività di ricerca e sviluppo, anche a livello internazionale, e, addirittura globale, in modo da evitare che qualche elemento critico venga trascurato.

Tuttavia, anche in presenza di una continuazione della crescita esponenziale della densità di integrazione, stanno acquistando crescente importanza dei fenomeni fisici, prima secondari, che riducono pesantemente i vantaggi marginali di questa evoluzione.

Si citano, nel seguito, i due fattori principali.

- **La crescita dei ritardi interni.** Siccome la di-

Una delle ragioni del successo dei transistor MOS è stata la bassa **potenza dissipata**, soprattutto in assenza di commutazioni. Per le tecnologie attuali questo non è più vero. Infatti, le potenze dissipate crescono rapidamente sia in funzionamento sia in *stand-by*. Per quanto riguarda il funzionamento, si ha un continuo aumento delle capacità parassite legate alle interconnessioni, che vanno fatte commutare in tempi sempre più brevi a causa dell'aumento della frequenza. Per quanto riguarda le condizioni di *stand-by*, le correnti parassite non sono più trascurabili, e la situazione, prevedibilmente, dovrebbe peggiorare: infatti, riducendo la tensione di alimentazione, occorre anche ridurre la tensione di soglia, per garantire correnti accettabili nel transistor acceso, e pilotare capacità sempre crescenti di interconnessione. Purtroppo, la fisica del funzionamento di un transistor MOS prevede che un transistor spento, con gate a massa, lasci passare una corrente residua che dipende esponenzialmente dalla tensione di soglia (Figura).

Indicativamente, la corrente parassita cresce di un ordine di grandezza per ogni ottanta-cento millivolt di riduzione della tensione di soglia. Riducendo le tensioni di alimentazione a 600 mV, come previsto per le tecnologie da 45 nm nel 2010, e le soglie a 100-200 mV, le correnti parassite in *stand-by* per dispositivi da decine di milioni di transistor, sarebbero dell'ordine di un Ampere!

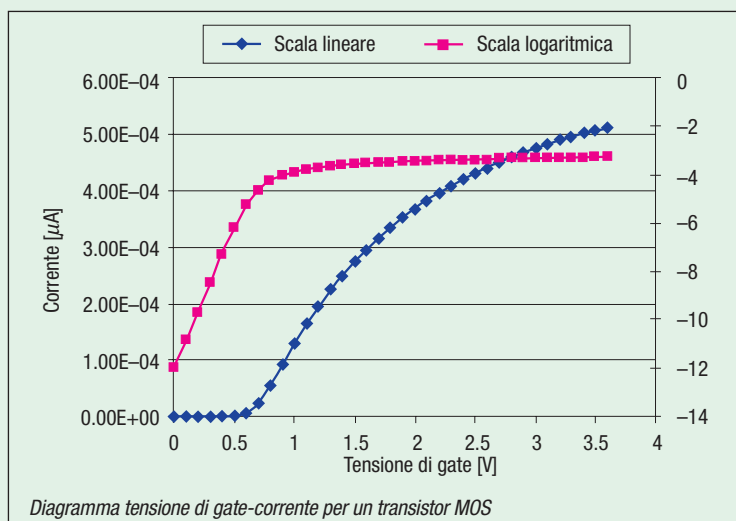


Diagramma tensione di gate-corrente per un transistor MOS

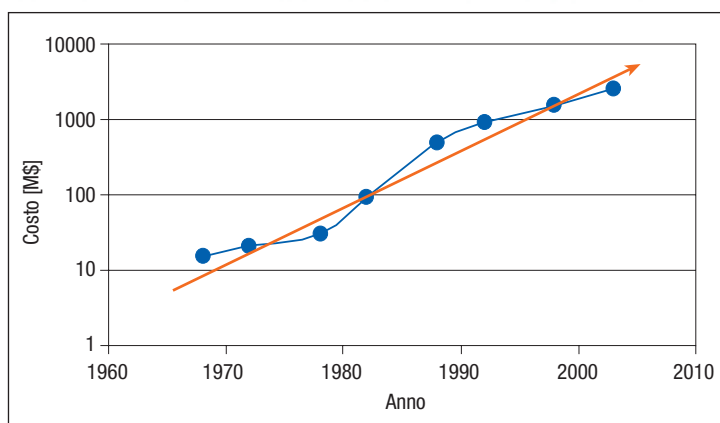
stanza tra le interconnessioni diminuisce, aumentano le capacità parassite; inoltre, a causa della diminuzione della sezione delle interconnessioni, aumentano le resistenze serie. Di conseguenza, diminuisce la velocità di propagazione dei segnali, proprio mentre si vuole aumentare la frequenza di funzionamento dei dispositivi.

□ **La crescita della dissipazione termica.** Con il passaggio a tecnologie scalate e nonostante la riduzione delle tensioni di alimentazione, le **potenze dissipate** dai circuiti CMOS, sia in funzionamento che in *stand-by*, sono in continua rapida crescita, tanto che i valori indicati dalla roadmap sono difficilmente compatibili con l'alimentazione a batterie, e addirittura con le tecnologie convenzionali di raffreddamento.

### 4.3. I limiti economici

Tutte le innovazioni tecnologiche e il miglioramento della qualità dei materiali che hanno reso possibile il processo di scalamento dei dispositivi hanno comportato, però, investimenti sempre crescenti in apparecchiature: da qui una seconda legge di Moore che si esprime affermando che, anche l'investimento necessario per sviluppare una nuova tecnologia, cresce in maniera esponenziale col tempo (Figura 4).

Al momento attuale, fermo restando il fatto che l'entità dell'investimento dipende in ma-



niera significativa dal tipo di prodotto in sviluppo e dalle economie di scala che si intendono effettuare, una stima intorno ai 2-5 miliardi di dollari non sembra lontana dal vero. Ogni nuova linea pilota richiede, quindi, investimenti (e coinvolge ricercatori) paragonabili con quelli degli acceleratori di particelle o dell'esplorazione spaziale. Anche se l'industria microelettronica spende tradizionalmente circa il 20% del proprio fatturato in nuove fabbriche e il 12-15% in ricerca e sviluppo, la crescita degli investimenti richiesti per una nuova linea pilota tende a rappresentare una porzione, sempre più alta, del fatturato, con alcune implicazioni economiche:

■ riduzione nel numero di società che si possono permettere linee pilota avanzate;

FIGURA 4

Seconda legge di Moore, sull'investimento necessario alla costruzione di una linea di produzione di circuiti integrati

I fenomeni di associazione di società diverse per condurre la ricerca in comune (SEMATECH, associazione di Motorola, Philips ed ST per la nuova linea da 300 mm ecc.);

I crescita dei rischi connessi ad un investimento sbagliato, che colpisce, soprattutto, le società che sviluppano le attrezzature di produzione nel settore della microelettronica.

In generale, si sta quindi assistendo ad un fenomeno di netto consolidamento del settore, sintomo di una industria matura, con alte barriere di ingresso, ed una forte riduzione della propensione a correre rischi. Si spiega anche così il forte successo della ITRS roadmap: rappresenta un tentativo di condividere il rischio, concordando un percorso di massima probabilità.

Un altro fattore limitante è dato dalla necessità di garantire un ritorno economico adeguato per gli investimenti fatti. Ogni nuova generazione tecnologica deve produrre utili sufficienti a ripagare le spese di sviluppo, e questo è possibile solo se si aprono nuovi mercati di massa. Potrebbe arrivare un momento in cui non esisteranno più applicazioni di massa tali da giustificare economicamente lo sviluppo di tecnologie a prestazioni superiori.

FIGURA 5

*Design gap: divario crescente tra le densità di transistor raggiungibili grazie alla tecnologia e quelle che possono essere gestite dalla progettazione*

## 5. QUALE FUTURO?

Che il processo di diminuzione esponenziale delle dimensioni del transistor MOS debba prima o poi finire è fuori discussione, soprattutto tenendo conto che già ora ci si sta avvicinando ad alcuni limiti fisici fondamentali, come il passo del reticolo cristallino e la solu-

bilità solida dei droganti nel silicio. D'altra parte la stessa sopravvivenza dell'industria elettronica, come la conosciamo ora, dipende in modo critico dalla continuazione della presente crescita esponenziale. Esiste quindi una forte spinta a cercare vie alternative che garantiscano la continuazione della situazione attuale per il più lungo tempo possibile. A questo proposito bisogna tenere presente che il punto critico non è costituito dalle dimensioni del transistor, ma *dalla possibilità di produrre circuiti integrati sempre più complessi a costi sempre più bassi*.

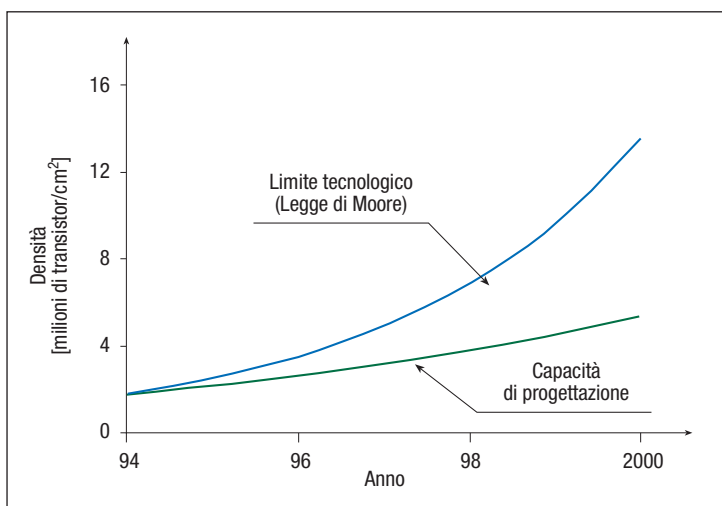
## 6. IL BREVE E MEDIO TERMINE

### 6.1. Soluzioni architetturali

Nel breve termine, due filoni di sviluppo che possono estendere le capacità di integrazione e le prestazioni della microelettronica, a costi relativamente contenuti, e senza coinvolgere rivoluzioni tecnologiche sono:

- a. l'aumento della efficienza di progettazione;
- b. l'integrazione su *package*.

Per quanto riguarda il primo fattore, bisogna riconoscere che l'aumento della complessità di progettazione, con il crescere del numero dei componenti, ha portato a soluzioni progettuali sempre meno efficienti, privilegiando il tempo di realizzazione piuttosto che la minimizzazione dell'area. Ciò è visibile nella curva della legge di Moore, dalle diverse velocità di crescita della densità di integrazione per memorie (strutture semplici e ripetitive) e per microprocessori (strutture complesse). Si è giunti, quindi, a parlare di *design gap* (Figura 5) per indicare il divario tra la crescita nel numero di dispositivi integrabili tecnologicamente e quelli realizzabili, in pratica, dai progettisti. Come si vede, il divario è notevole, ed un progresso significativo nelle metodologie di progettazione automatica potrebbe accrescere di un ordine di grandezza la densità di integrazione, senza richiedere nuove tecnologie. Va a questo punto sottolineato che, già da ora, sono note tecniche di progettazione che permetterebbero di incrementare notevolmente la densità dei circuiti integrati, riducendo allo stesso tempo la potenza dissipata, ma queste non vengono utilizzate per la mancanza di strumenti software di progettazione automatica, che consenta-







no di realizzare in tempi finiti circuiti della complessità desiderata.

Per quanto riguarda il secondo fattore, si sta cominciando a diffondere l'uso di assemblare diversi dispositivi integrati nello stesso package, per ridurre l'ingombro complessivo, per combinare dispositivi realizzati con tecnologie diverse, senza ricorrere a tecnologie ibride più costose e per aumentare le velocità di comunicazione tra i vari dispositivi, evitando di passare per la piastra stampata. Nuove tecnologie di assemblaggio stanno emergendo, finalizzate a questo scopo.

## 6.2. La fotonica

Sempre nell'ambito dell'integrazione su package, vale la pena di considerare le possibilità di utilizzare connessioni ottiche su chip. La spinta in questa direzione viene dai limiti che le interconnessioni tradizionali presentano al crescere delle frequenze di funzionamento, sia per il crescere dei ritardi legati all'aumento delle capacità parassite di interconnessione che per il peso crescente dei disturbi elettromagnetici.

L'idea ha guadagnato credibilità recentemente con la scoperta di tecnologie capaci di conferire al silicio (attraverso drogaggio con terre rare o la formazione di zone di silicio poroso) proprietà ottiche interessanti che, per quanto ben lontane da quelle dei semiconduttori composti, come l'arseniuro di gallio GaAs ed il fosforo di indio InP, rendono tuttavia possibile l'integrazione delle funzioni ottiche sul circuito integrato.

Rimane tuttavia il problema fondamentale che le dimensioni fisiche delle interconnessioni ottiche non possono essere inferiori a quelle della lunghezza d'onda utilizzata, mentre le interconnessioni elettriche hanno, già adesso, dimensioni molto inferiori. A questo va aggiunto il problema dei costi e della complessità tecnologica.

In questo contesto è quindi ipotizzabile che le interconnessioni ottiche su chip vengano limitate ad alcuni segnali, come quello di sincronizzazione interna (o *clock*) che devono essere distribuiti simultaneamente, per quanto possibile, ad un rilevante numero di dispositivi, e che già ora richiedono un livello di interconnessione dedicato, mentre l'uso

prevalente sarà per le interconnessioni tra chip diversi e su piastra stampata (*optical back-plane*) [12]. In questo quadro è possibile che il silicio finisca per occupare anche alcune delle nicchie sinora riservate ai semiconduttori composti.

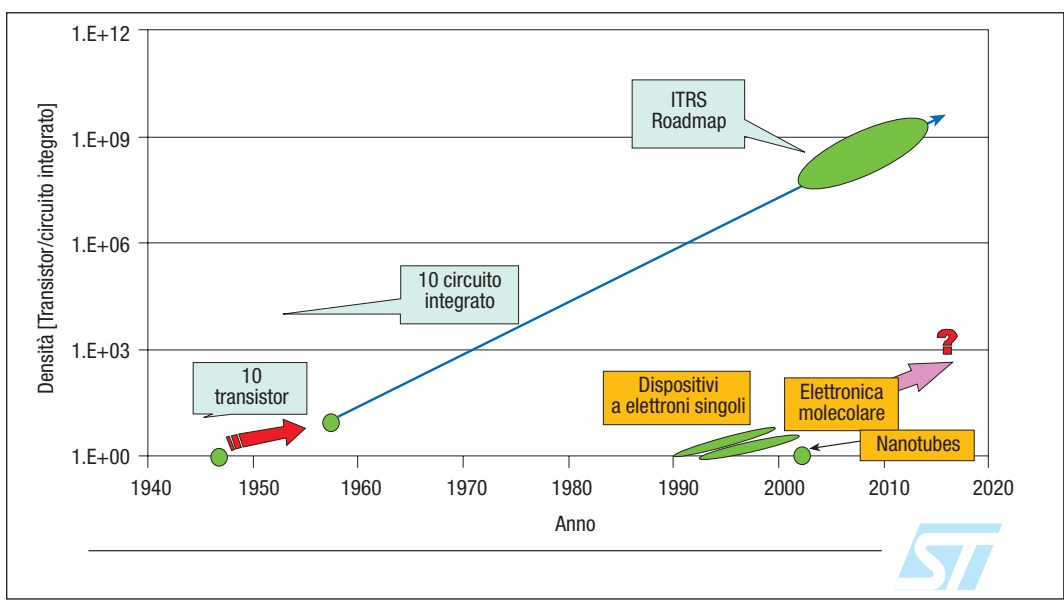
Dal punto di vista fisico, un segnale elettronico richiede energie di poco superiori al rumore termico (26 meV, a temperatura ambiente) per essere riconosciuto dal fondo, ma l'informazione si disperde su di una distanza di poche decine di nanometri; per contro, un fotone richiede energie di eccitazione molto più elevate (dell'ordine di 1 eV), ma si conserva per distanze molto maggiori. Risulta pertanto conveniente utilizzare la via ottica per il trasporto dell'informazione su lunghe distanze, mentre non esistono alternative all'elettronica per l'elaborazione del segnale e il suo trasporto a distanze inferiori.

## 7. IL LUNGO TERMINE

Iniziative come quelle finora considerate appartengono tutte ad uno schema evolutivo volto a migliorare ed estendere prestazioni e funzioni dei circuiti integrati, sempre basati sulla tecnologia MOS. Sul lungo periodo, però l'approccio evolutivo è necessariamente destinato a fallire, e occorrerà sviluppare tecnologie alternative.

Queste tecnologie, in genere conglobate sotto il termine nanoelettronica, sono l'oggetto del più grosso sforzo di ricerca accademica attualmente in corso, e progressi nel campo vengono annunciati pressoché tutti i giorni. Non va però dimenticato che l'obiettivo finale è di rimpiazzare la tecnologia CMOS quando questa giungerà al termine della sua evoluzione, fra qualche decina di anni, quando le complessità dei circuiti integrati saranno dell'ordine del miliardo di transistor. Un confronto delle scale temporali e di complessità è riportato in figura 6, e può dare un'idea delle dimensioni del problema. Quello che in genere si tende ad ignorare, è che il problema non è costituito dalla realizzazione di un transistor più piccolo. Già ora i circuiti integrati comprendono un livello di transistor e 6-8 livelli di interconnessioni, e una gran parte della potenza dissipata viene spesa per caricare e scaricare tutte le capacità parassite di interconnessione. Una

**FIGURA 6**  
 Percorso di crescita temporale della tecnologia Mos, la cui evoluzione giungerà al termine fra qualche decina di anni



nuova tecnologia dovrà in primo luogo risolvere il problema delle interconnessioni tra gli elementi attivi, se vorrà offrire una valida alternativa alla tecnologia MOS.

**7.1. Elettronica molecolare e nanotecnologie. Un'altra rivoluzione?**

Dato che sono già note singole molecole (con dimensioni dell'ordine di 3 nm) capaci di realizzare le funzioni tipiche dell'elettronica (rettificazione, memoria ecc.) [13], è ipotizzabile il loro utilizzo come dispositivi dal cui assemblaggio nascerà il circuito integrato.

Tra le varie molecole disponibili capaci di manipolare dati, un ruolo particolare è rivestito dalle molecole <sup>TM</sup> coniugate. Le molecole di questa classe (che comprende polimeri, fullereni e nanotubi) hanno la caratteristica essenziale di essere drogabili in maniera simile a quello che avviene col silicio, così da rendere possibile la traduzione dell'elettronica a base di silicio al mondo del carbonio. Per quanto la traduzione non sia completa (per esempio, in luogo del MOSFET si preferisce la struttura in cui il metallo è in diretto contatto con il canale formando una giunzione metallo-semiconduttore) e le prestazioni intrinseche siano di svariati ordini di grandezza peggiori di quelle dei corrispondenti dispositivi al silicio, una tale "elettronica organica" è una buona palestra per la soluzione di problemi di interfacciabilità, processabilità ecc. che l'elettronica molecolare pone. È inoltre

da notare che tra tutte le molecole di questa famiglia, la sintesi relativamente facile, l'alta stabilità ambientale e le prestazioni intrinseche dei nanotubi suggeriscono queste molecole come le candidate più interessanti dal punto di vista applicativo.

Il loro utilizzo pratico per la realizzazione di funzioni logiche complesse è però subordinato alla soluzione di due problemi fondamentali. In primo luogo, è necessario lo sviluppo di una tecnologia "dal basso verso l'alto" (*bottom-up*) e cioè capace di disporre le molecole-dispositivo in maniera ordinata ed accessibile con tecniche autoassemblanti (e cioè tali da non richiedere la disposizione delle molecole singolarmente, come è fin d'ora possibile utilizzando strumenti come il microscopio tunnel a scansione). Inoltre tale assemblaggio non può essere condizionato dalle caratteristiche intrinseche delle molecole, ma deve corrispondere ad un ordine arbitrario imposto dall'esterno. Il problema è quindi di tradurre il progetto del circuito integrato, costituito da milioni di porte logiche, o forse centinaia di milioni, considerando la scala temporale, in una disposizione corrispondente delle strutture molecolari, senza far uso delle costose tecniche litografiche, che costituiscono uno dei limiti delle tecniche CMOS convenzionali. Il secondo problema è costituito dall'estrazione a livello macroscopico (e quindi accessibile all'uomo) dell'informazione contenuta negli stati microscopici di atomi o molecole



che richiede, secondo l'analisi di Bohr, processi di amplificazione capaci di far evolvere l'apparato di misura su uno dei suoi possibili stati macroscopici in corrispondenza biunivoca con gli stati del sistema microscopico sotto osservazione [14]. Alla luce di ciò, e considerato il fatto che famiglie di dispositivi al silicio sono già capaci di comportarsi da rivelatori dello stato microscopico del sistema (diodi pin per la rivelazione di fotoni singoli, condensatori di memorie dinamiche per la rivelazione di particelle  $K$  ecc.), l'utilizzo più logico delle molecole-dispositivo è attraverso il loro autoassemblaggio su circuiti microelettronici. In questo modo si affiancherebbero le funzioni di manipolazione dei dati, realizzate con molecole-dispositivo, alle funzioni di amplificazione, indirizzamento e alimentazione, realizzate con dispositivi microelettronici. Questa architettura ibrida micro-nanoelettronica dovrebbe coniugare le tecnologie di miniaturizzazione (*top-down*) proprie della microelettronica a quelle di autoassemblaggio (*bottom-up*) proprie della nascente nanoelettronica [15].

## 7.2. Calcolo quantistico L'ultima rivoluzione?

La conducibilità di una singola molecola è usualmente controllata dallo stato di ossidazione di un singolo atomo. In ultima analisi, misurare la conduzione lungo una molecola determina la presenza o meno di un singolo elettrone. Nei dispositivi microelettronici convenzionali (come nei transistor delle memorie programmabili elettricamente) il passaggio dallo stato *on* a quello *off* richiede invece una carica dell'ordine di  $10^5$  elettroni. Peraltro, negli ultimi anni sono stati sviluppati dispositivi (SET, *Single-Electron Transistor*) il cui stato di conduzione è determinato dalla presenza o meno di un singolo elettrone. Al momento, questa classe di dispositivi richiede l'utilizzo di cluster di silicio di dimensioni nanometriche ("punti quantici") dispersi all'interno dell'ossido come siti ospitanti l'elettrone e gli effetti sono osservabili solo in condizioni criogeniche, il che ne rende problematico l'impiego nella pratica. Essi mostrano però come sia possibile arrivare alla costruzione *top-down* di dispositivi sensibili ad eventi microscopici singoli, e la loro applicazione a pic-

coli blocchi di memoria è stata dimostrata già da parecchi anni [16]. Tuttavia le tecnologie utilizzate sono ancora quelle della microelettronica convenzionale, spinte ai loro limiti fisici. Nel momento in cui si è reso accessibile alla tecnologia "di massa" il mondo microscopico, nasce spontaneo l'interesse verso un suo possibile utilizzo al fine di realizzare un'eventuale macchina di Turing quantistica (e cioè un calcolatore quantistico), in cui si sfrutta la proprietà fondamentale di ogni sistema microscopico per cui lo stato del sistema è la combinazione lineare di tutti i suoi possibili stati logici.

L'interesse di principio per il calcolo quantistico nasce dalla congettura di Feynmann per il quale la descrizione dell'evoluzione di un sistema dinamico quantistico richiede l'utilizzo di algoritmi di enorme complessità (si parla di "complessità esponenziale"). Assunta come vera questa congettura, essa può essere rovesciata nell'affermazione che la dinamica di un sistema quantistico contiene, seppure in maniera analogica, soluzioni di problemi di complessità esponenziale e quindi praticamente insolubili con macchine di Turing classiche. Successivamente (ma solo negli ultimi anni), verificata la congettura di Feynmann, si è dimostrata la possibilità concettuale di una macchina di Turing quantistica e si è trovato un problema che è di complessità esponenziale se trattato con macchine di Turing classiche, ma che diviene molto più semplice (si parla di "complessità polinomiale") se trattato con macchine di Turing quantistiche.

In seguito a questi risultati, si è sviluppato un ricco filone di ricerca che al momento sfiora soltanto il mondo dei circuiti integrati. La possibilità di manipolare e testare stati a elettrone (e fotone) singolo rende però plausibile che in un futuro più o meno lontano (ma al momento imprevedibile) il mondo del calcolo quantistico e quello della micro-nanoelettronica si fondano per una nuova rivoluzione.

## 8. CONCLUSIONI

L'evoluzione della microelettronica è stata caratterizzata da uno sviluppo esponenziale indipendentemente dalla particolare grandezza (densità di bit, densità di transistor, geometria minima) utilizzata per parametrizzare tale sviluppo.

Tale sviluppo, sostenuto da fattori tecnici ed economici e reso possibile dalla lontananza di limiti fisici fondamentali, è avvenuto su base spontanea per oltre 30 anni.

Le difficoltà progressivamente crescenti cui si va incontro nel processo di miniaturizzazione suggeriscono che tale sviluppo spontaneo è destinato a cessare in tempi brevi.

Al fine di non perdere i vantaggi che lo sviluppo esponenziale della microelettronica offre non solo al settore ma all'intera economia mondiale, si è passati da una crescita spontanea ad una assistita, attraverso una chiara identificazione dei problemi che occorre risolvere, perché la legge di Moore continui a valere.

È presumibile che la focalizzazione della ricerca sui colli di bottiglia allo sviluppo, e la concertazione delle attività da parte di diversi operatori (attraverso *joint venture*), riusciranno a sostenere la tendenza attuale fino a oltre il 2010; che la tendenza alla miniaturizzazione possa continuare oltre il 2015 o che lo sviluppo dell'elettronica passi attraverso nuovi tracciati tecnologici (come l'autoassemblaggio di dispositivi molecolari) resta argomento di intensa discussione.

### Bibliografia

- [1] Devoto G, Oli GC: *Vocabolario della Lingua Italiana*. Le Monnier, Firenze, 1979.
- [2] Moore GE: Progress in digital integrated electronics. *Proc. IEDM*, 1975.
- [3] Dennard RH: Design of Ion Implanted MOSFET's with very small dimensions. *IEEE-JSSC*, SC-9, 1974, p. 256.
- [4] Packan P: Pushing the limits. *Science*, Vol. 285, 1999, p. 2079.
- [5] Keyes RW: Fundamental limits of silicon technology. *Proc. IEEE*, Vol. 89, 2001, p. 223.
- [6] Frank DJ, Dennard RH, Nowak E, Solomon PM, Taur Y, Wong H-S P, Device scaling limits of Si MOSFETs and their application dependencies. *Proc. IEEE*, Vol. 89, 2001, p. 259.
- [7] Ito T, Okazaki S: Pushing the limits of lithography. *Nature*, Vol. 406, 2000, p. 1027.
- [8] Harriott LR. Limits of lithography. *Proc. IEEE*, Vol. 89, 2001, p. 366.
- [9] Lloyd S: Ultimate physical limits to computation. *Nature*, Vol. 406, 2000, p. 1047.

- [10] Ponomarev YV, Loo JJGP, Dachs CJ, Cubaynes FN, Verheijen MA, Kaiser M, van Berkum JGM, Kubicek S, Bolk J, Rovers M: *A Manufacturable 25 nm Planar MOSFET Technology*. Techn. Digest of 2001 Symp. on VLSI Technology, 33, 2001.
- [11] Gargini P, Doering RR: The 2001 ITRS from altitude. *Solid St. Technol.*, Vol. 46 n. 1, 2002, p. 72.
- [12] Meindl JD: *The Evolution of Monolithic and Polylithic Interconnect Technology*. Techn. Digest of 2002 Symp. on VLSI Technology, 2, 2002.
- [13] Joachim C, Gimzewski JC, Aviram A: Electronics using hybrid-molecular and monomolecular devices. *Nature*, Vol. 408, 2000, p. 541.
- [14] Bohr N: *The quantum postulate and the recent development of atomic theory, reprinted in N. Bohr, Atomic Theory and the Description of Nature*. (Cambridge Univ. Press, Cambridge, 1934).
- [15] Cerofolini GF, Ferla G: Toward a Hybrid Micro-Nanoelectronics. *J. Nanoparticle Res.*, (in stampa).
- [16] Yano K, Ishii T, Sano T, Mine T, Murai F, Seki K: *Single-Electron-Memory Integrated Circuit for Giga-To Tera Bit Storage*. Abstracts of Silicon Nanoelectronics Workshop, 1997, 22.

LIVIO BALDI laureato in Ingegneria Elettronica all'Università di Pavia nel 1973. Nel 1974 è entrato in SGS-ATES (ora STMicroelectronics) nel 1974, presso la Ricerca e Sviluppo Centrale, occupandosi dello sviluppo di tecnologie MOS e CMOS, e dell'integrazione di memorie non volatili. Ha poi coperto l'incarico di responsabile del gruppo di sviluppo delle piattaforme di progettazione per memorie non volatili, ed ora svolge il ruolo di coordinatore per i progetti di ricerca finanziati europei. Ha partecipato alla definizione dei programmi di lavoro di ESPRIT e di JESSI, e ha rappresentato la STMicroelectronics in molti progetti di ricerca europei. È autore di 18 brevetti USA e circa 45 tra pubblicazioni e presentazioni a conferenze.  
e-mail: livio.baldi@st.com

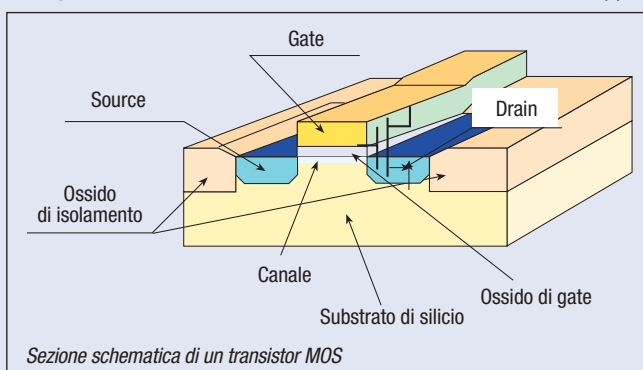
GIANFRANCO CEROFOLINI ottenuta la laurea in fisica all'Università di Milano nel 1970, inizia la sua carriera scientifica come ricercatore prima in Telettra (fino al 1977) e quindi in SGS (fino al 1988). Si sposta poi in EniChem dove diviene responsabile del Dipartimento Materiali Funzionali (fino al 1997) e torna infine in STMicroelectronics dove ricopre l'incarico di direttore scientifico della divisione Discrete and Standard Group. Ha svolto attività didattica come professore a contratto con le Università di Pisa e Modena, e con il Politecnico di Milano. È autore di circa 200 pubblicazioni su riviste internazionali. È inoltre autore (assieme a L. Meda) di un libro su Physical Chemistry of, in and on Silicon e curatore (con R. M. Mininni e P. Schwarz) di un libro su Chemistry for Innovative Materials.  
e-mail: gianfranco.cerofolini@st.com

## La Tecnologia MOS

I transistor di tipo MOS (abbreviazione di MOSFET, *Metal Oxide Semiconductor Field Effect Transistor*) funzionano in base all'opportuna giustapposizione di elementi semiconduttori, conduttori ed isolanti. Le caratteristiche degli elementi semiconduttori (tipicamente, silicio) vengono alterate grazie al loro "drogaggio" cioè all'iniezione di "impurità" con caratteristiche elettriche opportune. Quando il drogaggio fornisce caratteristiche conduttive dovute alla presenza di cariche negative (elettroni) in eccesso, allora il materiale risultante viene detto di tipo "N". Quando il drogaggio fornisce caratteristiche conduttive equivalenti alla presenza in eccesso di cariche positive (comunemente chiamate "lacune", cioè mancanze di elettroni, anche se queste non hanno per riscontro un elemento fisico reale), allora il materiale risultante viene detto di tipo "P". Mettendo in contatto semiconduttori drogati P e N, si forma una giunzione rettificante, o diodo, cioè una struttura che consente il passaggio di corrente se si applica la tensione in una direzione (polarizzazione diretta), mentre presenta resistenza elevatissima se la tensione è applicata nella direzione opposta (polarizzazione inversa).

### FUNZIONAMENTO DEL TRANSISTOR NMOS

È composto da tre elettrodi. Due, il *source* e il *drain* sono costruiti in materiale semiconduttore fortemente drogato di tipo "N", impiantati su un substrato semiconduttore di tipo "P" drogato molto debolmente. Le zone fortemente drogate (chiamate source e drain) sono in contatto *ohmico* con opportune metallizzazioni, per cui è possibile stabilire una differenza di potenziale tra le due (ed anche rispetto al substrato, detto *body*, che anch'esso ha un contatto metallico). Il terzo elettrodo, il *gate*, sormonta il substrato di tipo "P", ne è separato da un sottile strato isolante (ossido di gate) ed è costruito in modo da sovrastare la zona di substrato che separa il source dal drain (zona di *canale*). Uno schema del dispositivo è riportato nella figura. Se il gate è tenuto alla stessa tensione del substrato, anche se si applica una tensione tra le zone di source e di drain, non si ha passaggio di corrente, in quanto sono separate da una regione di drogaggio opposto, e quindi si comportano come diodi polarizzati in inversa. Al contrario, quando il gate viene portato ad una tensione positiva superiore ad un valore critico, detto *tensione di soglia*, (dell'ordine delle centinaia di millivolt), gli elettroni vengono attirati alla superficie formando un percorso conduttivo tra il source ed il drain (*canale*), tra i quali può quindi fluire una corrente elettrica. Si realizza, così, un interruttore elettrico il cui stato, aperto o chiuso (non conduttivo, conduttivo) è controllato elettricamente (dalla tensione sul suo elettrodo di gate). Le prestazioni di un transistor MOS, cioè la corrente che è in grado di far passare ad un certo valore delle tensioni applicate, sono determinate dalla distanza tra source e drain, definita *lunghezza di canale* e dallo spessore del dielettrico di gate. Per un insieme di motivi, essenzialmente legati alla buona stabilità chimica, alla facilità di realizzazione e alla impermeabilità ai contaminanti più comuni, la scelta del migliore dielettrico si è rivelata quella del  $\text{SiO}_2$  cresciuto per ossidazione ad alta temperatura del silicio, il che ha contribuito a imporre il silicio come semiconduttore base per la microelettronica.



### FUNZIONAMENTO DEL TRANSISTOR PMOS

È del tutto analogo a quello del transistor NMOS, se non che le polarità delle tensioni ed il tipo di droganti sono invertiti (drain e source fortemente drogati con impurità di tipo "P" costruiti su un substrato di tipo "N"). La corrente tra il source e il drain può fluire quando sul gate si ha una bassa tensione, mentre è interdetta quando la tensione di gate è alta.

### L'ARCHITETTURA CMOS

Avendo a disposizione transistor MOS di due tipi diversi, P o N, le cui caratteristiche sono complementari, è possibile collegarli in serie nello stesso dispositivo per realizzare le funzioni logiche, con caratteristiche molto interessanti. Si consideri, ad esempio, la struttura logica più elementare, l'inverter CMOS (*complementary MOS*), rappresentato in sezione schematica in figura A), e come schema elettrico in figura B). Essendo le due gate connesse tra di loro e al segnale di ingresso, quando questo è a livello alto, il transistor di tipo N sarà in conduzione, e avrà una bassa resistenza serie, ma quello di tipo P sarà spento, e presenterà un'altissima resistenza, e quindi l'uscita sarà bassa. Il comportamento sarà l'opposto quando l'ingresso sarà basso. Il circuito si comporta, quindi, come una serie di due interruttori, schematizzata in figura C), e si ha passaggio di corrente solo durante le transizioni di stato logico. Una funzione logica (in questo caso, la negazione) realizzata in CMOS ha quindi la proprietà di non dissipare potenza se non durante la commutazione.

