

IL MONDO DEI *CHIP*: PROGRESSI TECNOLOGICI E VULNERABILITÀ DI UNA FILIERA STRATEGICA

Nota del s.c. ANDREA L. LACAITA (*)

(Adunanza del 24 novembre 2022)

SUNTO. – Dopo una breve descrizione degli sviluppi che hanno portato alla nascita della microelettronica, la nota illustra le leggi economiche e fisiche che da oltre cinquant'anni guidano la crescita di questo segmento tecnologico. Seguendo le leggi di Moore e Dennard, la complessità dei *chip* microelettronici è aumentata vertiginosamente, giungendo a incorporare componenti sempre più miniaturizzati, con dimensioni ormai nanometriche e in numero paragonabile alle stelle di una galassia. Tuttavia, nel corso di questi decenni la microelettronica non ha perseguito soltanto la strada della miniaturizzazione. Accanto alle tecnologie digitali si sono sviluppate piattaforme diversificate che permettono di realizzare *chip* altrettanto essenziali nei segmenti di prodotto non digitali. Gli attori di tutte queste filiere produttive sono distribuiti sui vari continenti e sono fortemente interconnessi. Mentre gran parte della produzione avanzata di *chip* digitali è in Asia, la maggior parte della loro progettazione risiede negli Stati Uniti mentre le aziende europee sono all'avanguardia nei *chip* per la sensoristica e per le applicazioni industriali e automobilistiche. Tuttavia, la recente crisi pandemica e le crescenti tensioni geopolitiche hanno evidenziato come la mancata fornitura di un solo *chip* possa interrompere la produzione di beni essenziali. Pertanto, i governi delle economie avanzate stanno attuando iniziative per ricollocare logisticamente i siti di fabbricazione e per rafforzare le capacità tecnologiche locali con l'obiettivo di proteggere il tessuto economico da crisi internazionali. Anche l'Europa si sta muovendo in questa direzione ma, per essere efficace, è sempre più impellente risolvere alcune debolezze della costruzione politica europea che anche in questo campo limitano il ruolo del nostro continente sulla scena globale.

(*) Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Milano, Italy. E-mail: andrea.lacaita@polimi.it

ABSTRACT. – After a brief coverage of the developments that led to the birth of microelectronics, the note describes the economic and physical laws that have guided the growth of this technological segment for over fifty years. Following the laws of Moore and Dennard, the complexity of microelectronic chips has increased dramatically, reaching the point of incorporating miniaturized components, now nanometric in size, comparable in number to the stars in a galaxy. However, during these decades microelectronics has not only pursued the path of miniaturization. Alongside digital technologies, diversified platforms have developed that allow for the creation of equally essential chips in non-digital product segments. The actors of all these production chains are distributed over the various continents and are strongly interconnected. While much of the advanced digital chip manufacturing is in Asia, most of their design resides in the United States while European companies are at the forefront of sensor, industrial and automotive chips. However, the recent pandemic crisis and rising geopolitical tensions have highlighted how the lack of just one chip can disrupt the production of essential goods. Therefore, the governments of advanced economies are implementing initiatives to logistically relocate manufacturing sites and to strengthen local technological capabilities with the aim of protecting the economies from international crises. Europe too is moving in this direction but, in order to be effective, it is increasingly urgent to solve some weaknesses in the European political construction which also in this field limit the role of our continent on the global scene.

1. CENNI STORICI INTRODUTTIVI

I *chip* o circuiti integrati sono componenti miniaturizzati realizzati su strati di materiale semiconduttore, tipicamente silicio. Essi sono il prodotto della tecnologia microelettronica che prese le mosse dall'invenzione del transistor, avvenuta nel 1947 ad opera di J. Bardeen, W. H. Brattain e W. Shockley presso i laboratori della compagnia telefonica americana Bell. Per comprendere il contesto e la portata di questa invenzione e dello sviluppo tecnologico e industriale che ne seguì, è necessario fare qualche passo indietro.

Il triodo è un componente elettronico a tre terminali che ha un ruolo centrale nell'elaborazione dei segnali. Anzi, si può senz'altro affermare che l'elettronica, intesa come scienza e tecnologia dell'informazione, nasce proprio con l'invenzione del triodo ad opera di Lee De Forest nel 1906 (*Fig. 1*) [1]. Quando De Forest aggiunse alla valvola termoionica di Fleming il terzo terminale, la "griglia", aprì la strada alla realizzazione di due funzioni fondamentali: la commutazione e l'amplificazione.

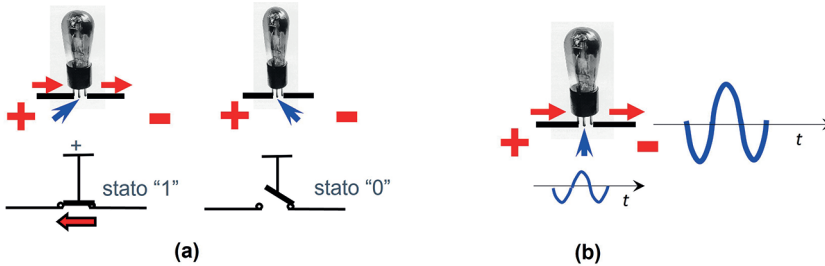


Fig. 1. Illustrazione schematica delle funzioni realizzate con un triodo a vuoto: (a) commutazione; (b) amplificazione.

Per commutazione si intende la possibilità di consentire o interrompere il passaggio di corrente tra i due terminali principali del triodo, agendo sul potenziale elettrico della griglia (Fig. 1a). Il triodo si comporta così come un interruttore comandato elettricamente. Ai due stati di conduzione si può associare lo stato di una variabile binaria, “0” – “circuito aperto” e “1” – “circuito chiuso”. Da questa associazione prende le mosse la realizzazione dei circuiti per il calcolo binario impiegati poi nella realizzazione dei primi calcolatori elettronici digitali (1939 Atanasoff-Berry Computer, 1946 ENIAC).

Il controllo della conduzione elettrica di un triodo abilita anche un'altra funzione fondamentale: l'amplificazione. In questo caso non si tratta di commutare tra i due stati estremi di conduzione, ma di variare con continuità il potenziale di controllo del terminale di griglia. Questa variazione induce un'analogia variazione della corrente nel componente (Fig. 1b). L'andamento nel tempo della corrente replica l'andamento del segnale di tensione applicato all'elettrodo di controllo. In questa disposizione il triodo si comporta come un rubinetto su un circuito idrico in cui il terminale di griglia opera da manopola. Nel rubinetto, una azione limitata sulla manopola permette di variare la posizione della saracinesca da questo pilotata. Nonostante la ridotta quantità di energia meccanica impiegata, questa azione può cambiare significativamente la portata di acqua che attraversa la condotta principale controllata dalla saracinesca e quindi l'energia meccanica che si trasferisce a valle del rubinetto. Si ha quindi una amplificazione: impiegando una piccola quantità di energia se ne permette l'erogazione di una quantità ben maggiore. Anche nel triodo la limitata energia elettrica necessaria a variare il potenziale dell'elettrodo di griglia può far cambiare in modo importante la corrente tra i terminali principali del componente e quindi l'energia elettrica erogata a valle.

Proprio l'amplificazione dei segnali elettrici, permessa dall'uso del triodo, consentì lo sviluppo della radio, della televisione, del radar, delle prime reti di telecomunicazione.

Tuttavia, la fragilità dei tubi a vuoto, le alte tensioni di funzionamento ne limitavano le applicazioni sia nelle trasmissioni che nelle prime macchine da calcolo. L'ENIAC, il primo calcolatore elettronico digitale di uso generale, era caratterizzato da 17468 tubi a vuoto, occupava un'area di ben 180 m² e dissipava una potenza di 174 kW. All'indomani della seconda guerra mondiale, l'atmosfera di rinnovata fiducia nel futuro portava a immaginare uno sviluppo significativo delle reti di telecomunicazione, che tuttavia richiedevano la sostituzione dei triodi con componenti più compatti e robusti. Questa necessità diede nuovo impulso alla ricerca sulle proprietà dei materiali semiconduttori, con l'obiettivo di realizzare un triodo a stato solido.

L'idea di un triodo a stato solido era già prefigurata nei brevetti di J. E. Lilienfeld (1928) [2] e O. Heil (1935) [3]. Si immaginava di variare la conducibilità di uno strato conduttivo sfruttando l'azione di un intenso campo elettrico generato da un elettrodo di controllo. Dopo diversi tentativi il transistor vide la luce il 23 dicembre 1947 ad opera di J. Bardeen, W. H. Brattain (*Fig. 2*) [4]. Successivamente, W. Shockley, che aveva contribuito a quelle ricerche, inventò il transistor bipolare a giunzione [5] che ben presto si affermò come il dispositivo con prestazioni migliori nella prima fase di diffusione dei nuovi componenti microelettronici. Nel 1956, per questi risultati, J. Bardeen, W. H. Brattain e W. Shockley ricevettero il Premio Nobel per la Fisica.

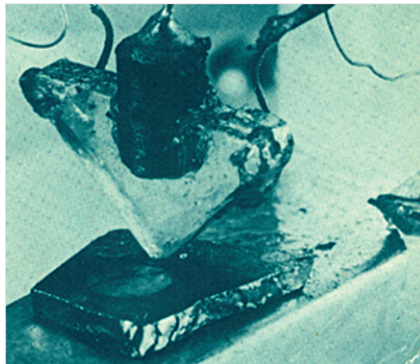


Fig. 2. Il transistor a punta di contatto.

Con l'invenzione del transistor inizia l'era dell'elettronica su scala miniaturizzata. Negli anni '50 i transistori cominciarono a sostituire i tubi a vuoto e furono alla base dello sviluppo di nuovi prodotti come le prime radio portatili. Ma per giungere alla microelettronica odierna fu necessario un altro passo importante: la realizzazione del circuito integrato o "*microchip*". Invece di realizzare un circuito elettronico connettendo i componenti prodotti singolarmente con conduttori elettrici, si puntò a costruire l'intero insieme di componenti direttamente su un substrato di semiconduttore, fabbricandoli contemporaneamente con gli stessi passi tecnologici. Questo approccio prometteva di ridurre il costo di fabbricazione realizzando, peraltro, un circuito più robusto e compatto. L'idea di realizzare sullo stesso strato di germanio più componenti fu maturata da Jack Kibly della Texas Instrument [6] e parallelamente da Jean Hoerni e Bob Noyce di Fairchild. Fu Jean Hoerni a gettare le basi tecnologiche del "processo planare" [7], permettendo a Fairchild di realizzare il primo processo planare di fabbricazione di un *microchip* sulla superficie di uno strato semiconduttore (1961). Tra i primi straordinari successi ottenuti con il processo planare c'è la realizzazione del primo microprocessore Intel 4004 (1971), progettato dall'italiano Federico Faggin [8]. Impiegava circa 2300 transistori, consumava appena 1W e occupava l'area di soli 12mm².

Il processo planare prende le mosse da substrati di silicio che sono disponibili in "fette" o *wafer* (Fig. 3). Si tratta di strati di forma circolare, con diametro variabile da circa 50 a 300mm e spessore dell'ordine di 0,5mm. Questi si presentano con una superficie frontale speculare. La Fig. 3 illustra schematicamente una semplice sequenza di lavorazione. In Fig. 3a la superficie del *wafer* viene dapprima ossidata esponendola ad ossigeno o vapore d'acqua ad alta temperatura. Successivamente su di essa si distribuisce un *resist* polimerico che, una volta essiccato, compie le funzioni di una pellicola fotografica. Tra una sorgente ottica e la superficie del *wafer* è interposta una "maschera", ovvero un substrato di quarzo su cui è deposto uno strato di cromo. La geometria che deve essere proiettata sulla superficie del *wafer* è definita rimuovendo lo strato di cromo solo in alcune aree specifiche e rendendo quindi selettivamente trasparente la maschera di quarzo alla luce generata dalla sorgente. Quando il *wafer* è irradiato, in corrispondenza delle aree non protette dalla maschera, il polimero è impressionato, cambia le sue proprietà chimiche divenendo, per esempio, solubile. Dopo l'esposizione, immergendo il *wafer* in una soluzione di svi-

luppo, il *resist* nelle regioni illuminate viene dissolto, lasciando la superficie del *wafer* esposta alle successive lavorazioni. Nelle altre zone, invece, la superficie resta protetta dal *resist*. In questo modo è possibile agire selettivamente per rimuovere, ad esempio, l'ossido nelle zone non protette come in Fig. 3*b*. Successivamente, nell'esempio in Fig. 3*b*, il *wafer* è esposto ad una impiantazione di ioni che modifica le proprietà elettriche del substrato nelle regioni esposte, ottenendo la struttura in Fig. 3*c*. Impiegando passi simili a quelli appena illustrati, si possono fabbricare contemporaneamente sulla superficie del *wafer* più repliche dello stesso microcircuito, disposte l'una accanto all'altra. Al termine della lavorazione, il *wafer* viene tagliato separando i singoli microcircuiti e ottenendo così i *chip*. Queste piastrine di silicio sono poi montate e incapsulate in contenitori di resina e montate su schede elettroniche.

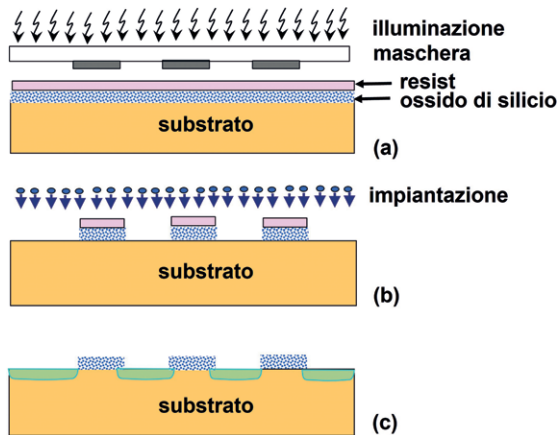


Fig. 3. Illustrazione schematica del processo di fabbricazione planare.

2. LA MICROELETTRONICA E LA MINIATURIZZAZIONE

L'illustrazione precedente deve essere intesa come una semplificazione di processi di fabbricazione molto sofisticati. La realizzazione di un *chip* richiede centinaia di passi con controlli accuratissimi delle condizioni a cui sono esposti i *wafer* (temperatura, atmosfere gassose, specie chimiche, sorgenti ottiche, tempi, etc..) e i relativi impianti sono estremamente costosi. Oggi gli investimenti di capitale necessari

per realizzare una fabbrica in grado di produrre *chip* nelle tecnologie digitali più avanzate superano i 10mld\$. E' evidente che per ammortizzare simili investimenti è indispensabile che ogni fabbrica generi decine di miliardi di componenti. L'analisi dei costi in un ambito così sofisticato è un compito complesso. Tuttavia per comprenderne alcuni aspetti è possibile ricorrere ad alcuni modelli semplificati. Si consideri, per esempio, il costo legato alla realizzazione di un *chip*. Questo costo può essere espresso come rapporto tra il costo necessario per lavorare un *wafer* di silicio e il numero di *chip* realizzati sulla superficie del *wafer*, ovvero

$$\$/chip = \frac{\$/wafer}{chip/wafer}. \quad (1)$$

Per ridurre il costo del *chip* è necessario ridurre il valore al numeratore. A questo scopo la fabbrica deve operare a pieno regime e nel corso della sua vita deve produrre decine di miliardi di *wafer*. In questo modo l'incidenza del capitale investito sul costo di ogni *wafer* si riduce progressivamente (Fig. 4) giungendo infine ad eguagliare il solo costo marginale determinato dal processo di fabbricazione. Solo operando su grandi volumi, è possibile abbattere l'incidenza degli investimenti di capitale sul costo. Le produzioni microelettroniche sono quindi intrinsecamente legate alla realizzazione di prodotti di ampia diffusione, ovvero pervasivi.

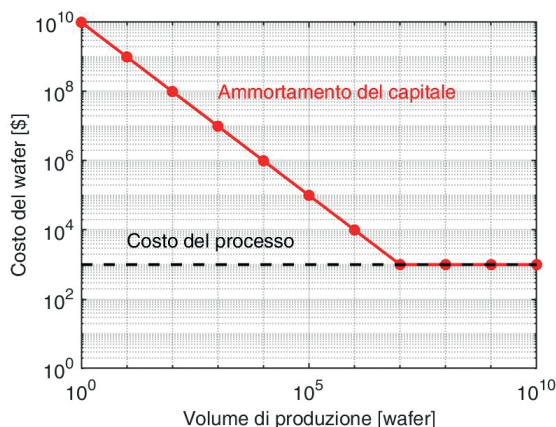


Fig. 4. Incidenza del capitale investito sul costo del *wafer* prodotto in funzione del volume di produzione. Per grandi volumi il costo del *wafer* è limitato dal costo del processo e non dall'ingente investimento iniziale.

Per ridurre ulteriormente il costo, si deve agire anche sul denominatore della frazione in (1) aumentando il numero di *chip* realizzati sull'area del *wafer*. Questo numero dipende dalla dimensione del *wafer*, che è infatti aumentata nel corso degli anni, giungendo oggi a diametri di 300mm, e dalla dimensione dei singoli transistori. Le dimensioni minime dei transistori dipendono dall'accuratezza con cui è possibile proiettare sulla superficie del *wafer* i contorni delle geometrie definite sulla maschera. Questa accuratezza, a sua volta, è limitata dal fenomeno della diffrazione ottica. Nelle produzioni dei circuiti digitali più avanzati, per ridurre l'effetto della diffrazione e per permettere una ulteriore miniaturizzazione delle dimensioni dei dispositivi si impiegano sorgenti di radiazione nel profondo ultravioletto (*Extreme Ultra Violet – EUV - Lithography*).

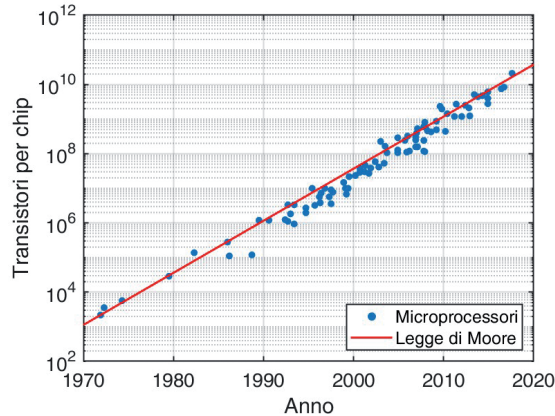


Fig. 5. Crescita del numero di transistori nei *chip* di microprocessori e andamento della legge di Moore assumendo che il raddoppio del numero di transistori di un *chip* avvenga in 24 mesi. Figura adattata da [11].

Pervasività e miniaturizzazione sono quindi tendenze strutturali delle produzioni microelettroniche, strettamente connesse alla sostenibilità economica degli ingenti investimenti richiesti dal settore. Già nel 1965 G. Moore, allora in Farchild e successivamente tra i fondatori di Intel, osservava come la riduzione dei costi spingeva alla miniaturizzazione dei componenti e al miglioramento delle rese di produzione, determinando il raddoppio del numero di transistori nei *chip* prodotti ogni 12 mesi. Questa osservazione induceva a prefigurare *chip* con 65000 transistori nel 1975 [9]. Corretta successivamente, portando il tempo di rad-

doppio del numero di transistori a 24 mesi [10], questa legge empirica, detta Legge di Moore, è diventata una linea guida dello sviluppo del settore nei successivi decenni e ancora oggi rappresenta correttamente il tasso di crescita del numero di transistori contenuti nei *chip* digitali più avanzati (Fig. 5) [11]. Oggi questo numero supera i 10^{10} transistori per *chip*, un valore dell'ordine del numero di stelle di una galassia.

Tuttavia la sola spinta economica alla miniaturizzazione non sarebbe stata sufficiente se essa non fosse stata accompagnata da un'altra caratteristica: diminuendo le dimensioni del transistore migliorano le prestazioni connesse alla commutazione e quindi alle applicazioni digitali. Questo aspetto chiave fu evidenziato da R. H. Dennard, ricercatore di IBM e inventore delle memorie DRAM. Le leggi della miniaturizzazione dei transistori, ricavate da Dennard, sono riportate nella tabella in Fig. 6 [12], assieme alla struttura schematica di un transistore MOSFET, dispositivo impiegato nelle applicazioni digitali. I terminali principali del componente sono indicati come (D) – (*Drain*) e (S) – (*Source*), mentre l'elettrodo di controllo, simile alla “griglia” del triodo a vuoto, è il terminale (G) – (*Gate*). Le prestazioni di questo transistore dipendono dal campo elettrico, F , che spinge le cariche tra i terminali principali (D) ed (S). Il suo valore medio può essere stimato dal rapporto

$$F \approx \frac{V_{DS}}{L}, \quad (2)$$

dove V_{DS} è la tensione tra i terminali (D) ed (S) e L è la loro distanza, detta anche *lunghezza di canale*. Si supponga ora di ridurre la distanza L di un fattore k . Per mantenere costante il valore di F , anche la tensione V_{DS} deve ridursi del fattore k . Se il campo elettrico resta costante, la velocità di moto delle cariche che si muovono tra i due terminali resta costante mentre il tempo necessario per attraversare il componente diminuisce. Infatti, scrivendo il tempo τ come rapporto tra la lunghezza L e la velocità media delle cariche, si ha:

$$\tau \approx \frac{L}{v}, \quad (3)$$

Se L diminuisce di un fattore k e v resta costante anche τ diminuisce di k . Il tempo τ non è solo il tempo di transito delle cariche tra i terminali del dispositivo ma anche il tempo necessario a riempire o svuotare di cariche il dispositivo. Esso, in definitiva, fissa il tempo necessario al dispositivo per commutare tra lo stato “0” e “1”. Quindi, se τ dimi-

nuisce di un fattore k il numero di commutazioni al secondo che il componente è in grado di fare aumenta come k e quindi aumenta come k il numero di operazioni digitali che si possono compiere nell'unità di tempo. La tabella in Fig. 6 riporta come variano anche le altre grandezze. Se tutte le dimensioni lineari del componente si riducono di k , non solo la lunghezza L ma anche la sua larghezza diminuisce di k . Quindi la corrente che lo attraversa diminuisce di un fattore k e l'area del componente diminuisce di un fattore k^2 . La potenza elettrica, ovvero il consumo di energia, che dipende dal prodotto tra la tensione e la corrente ai terminali del componente, diminuisce di un fattore k^2 . In definitiva, le relazioni ricavate da Dennard, dette anche leggi di *scaling* a campo costante, mettono in evidenza come la miniaturizzazione permetta non solo di ridurre l'area e il costo del componente, ma anche di compiere un maggior numero di calcoli nell'unità di tempo con un minor consumo di energia. La spinta alla miniaturizzazione, indicata da G. Moore come esigenza economica legata alla riduzione del costo del componente trova quindi nelle leggi di Dennard ulteriore e determinante motivazione divenendo lo strumento principale per migliorare la capacità di elaborazione digitale dei sistemi elettronici.

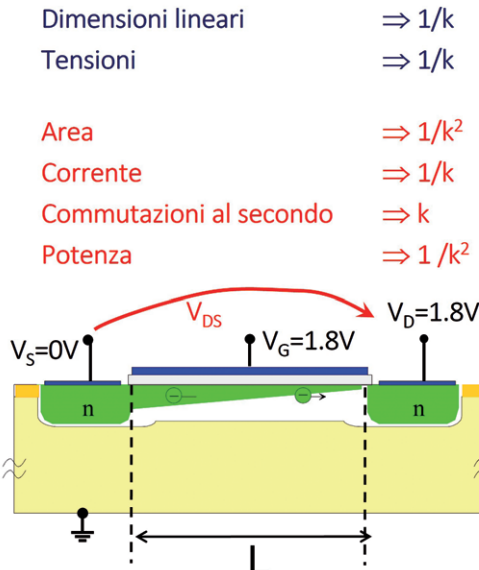


Fig. 6. Legge di *scaling* a campo elettrico costante e dipendenza delle prestazioni del transistor MOSFET al diminuire delle dimensioni lineari e delle tensioni applicate.

In questo contesto, dominato dalla corsa alla miniaturizzazione, fu quindi naturale incominciare a qualificare i processi tecnologici messi a punto dai vari produttori riferendosi alla minima dimensione che essi erano in grado di realizzare sul *wafer*. Si introduce quindi il concetto di *nodo tecnologico* espresso in nm (p.es. 180nm, 130nm, 90nm, 65nm, etc...). La Fig. 7 riporta l'andamento storico dei nodi tecnologici. Si noti come il rapporto tra due numeri contigui nella sequenza è circa $\sqrt{2}$, poiché $k=\sqrt{2}$ avrebbe dovuto essere il fattore di *scaling* da applicare alle dimensioni lineari dei componenti passando da una generazione alla successiva. Ciò avrebbe permesso, secondo le leggi di Dennard, la diminuzione di un fattore 2 dell'area del singolo transistor e il raddoppio del numero di transistori realizzabili per ogni cm^2 di *wafer*. In effetti, almeno inizialmente, questi numeri furono indicativi della minima dimensione, L , del canale dei transistori (Fig. 7). Successivamente, tra il 1995 e il 2005 la dimensione fisica del canale è diminuita anche più rapidamente del previsto per ottenere un miglioramento delle prestazioni. Poi il suo tasso di riduzione è rallentato per il sopraggiungere di limiti fisici. Un canale di 20nm corrisponde a solo 40 passi atomici tra i terminali principali del transistor. Oggi la lunghezza di canale si attesta su valori di poco inferiori a 20nm [13]. Tuttavia, agendo sulle altre dimensioni del componente e sulle distanze tra le interconnessioni il numero di transistori realizzati sui *chip* digitali continua ad aumentare e a seguire abbastanza bene la legge di Moore (Fig. 5).

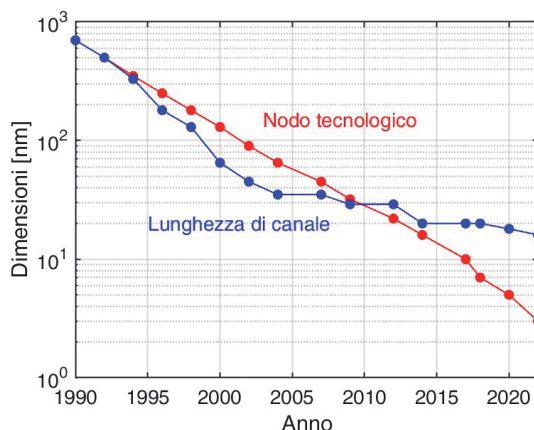


Fig. 7. Evoluzione dei nodi tecnologici e della distanza tra i terminali principali del transistor MOSFET.

3. LA DIVERSIFICAZIONE

Le tecnologie digitali si sono quindi sviluppate sotto la spinta della legge di Moore e seguendo le leggi di Dennard con gli opportuni aggiornamenti [14]. Diverso è stato invece il processo evolutivo delle tecnologie dedicate all'amplificazione dei segnali, alle comunicazioni e alle applicazioni di potenza. Se infatti la miniaturizzazione ha effetti benefici sulle prestazioni digitali poiché favorisce la commutazione rapida del transistor, questo non è vero per le proprietà di amplificazione. Al diminuire delle dimensioni del componente le massime tensioni applicabili diminuiscono come k . Diminuisce anche l'amplificazione di tensione ottenibile, così come la massima potenza erogabile ai dispositivi posti a valle. Di conseguenza, in molte applicazioni importanti come quelle industriali, in cui si ha spesso a che fare con l'azionamento di motori elettrici, o nelle comunicazioni in cui è necessario erogare potenza elettrica alle antenne, i transistori impiegati non sono di dimensione minima. In questi ambiti, per motivi fisici, non si può seguire il modello economico tracciato dalla legge di Moore e dalle leggi di *scaling* di Dennard. Questa divergenza è rappresentata pittoricamente in Fig. 8 dove, ortogonalmente alla direttrice dominata dalla miniaturizzazione e dai crescenti investimenti in infrastrutture produttive, si sviluppa una traiettoria di diversificazione che punta a promuovere soluzioni microelettroniche nei segmenti di prodotto non digitali.

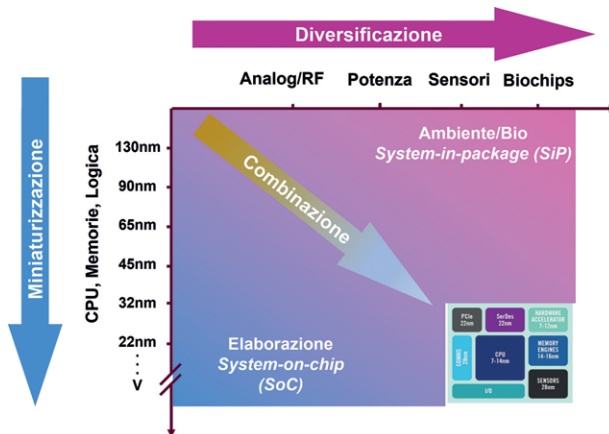


Fig. 8. Evoluzione delle piattaforme tecnologiche microelettroniche tra miniaturizzazione e diversificazione.

Un passo pionieristico in questa direzione fu compiuto da STMicroelectronics che verso la metà degli anni '80 sviluppò un processo dedicato alla realizzazione di componenti di potenza. Si trattava della tecnologia BCD, adottata ancora oggi, che permette di realizzare su *wafers* non solo i componenti classici dell'elettronica digitale (transistori MOSFET e memorie) ma anche transistori bipolari che garantiscono ottime prestazioni di amplificazione del segnale e transistori di potenza, in grado di sopportare alte tensioni. Questa tecnologia, riconosciuta come pietra miliare nello sviluppo dell'elettronica (*IEEE Milestone Award 2021* [15]) ha permesso all'azienda, l'allora italiana SGS, di proporre un solo processo integrato per sviluppare microsistemi per applicazioni all'automobile, agli elettrodomestici, ai dispositivi medicali, agli azionamenti di motori, schermi, modem, dischi rigidi etc..

Nei settori non digitali la competizione tra i diversi attori industriali non è basata sulla rapidità con cui si riducono le dimensioni dei singoli componenti (memorie e processori) per guadagnare margini dal minore costo di fabbricazione rispetto alla generazione precedente. La competizione si fonda invece sulla identificazione di un processo microelettronico che, impiegando nuovi circuiti, dispositivi e materiali, possa offrire migliori prestazioni o aprire applicazioni completamente nuove. Si pensi per esempio alla realizzazione in silicio di sensori micro-meccanici che possano essere in grado di rilevare le accelerazioni del corpo a cui il *chip* è connesso. Queste tecnologie, che non necessitano di miniaturizzazione spinta dei componenti, hanno aperto nuove applicazioni nei settori della sicurezza automobilistica con la realizzazione dei sensori integrati per gli *airbag*, o dei videogiochi con la realizzazione di interfacce che permettono all'utente di interagire con la realtà virtuale. Si tratta quindi di comparti a minore intensità di capitali investiti, perché non è necessario competere sulla miniaturizzazione spinta del singolo transistor, ma con una maggiore creatività progettuale e varietà tecnologica in termini di materiali e processi, in cui possono operare aziende con capacità finanziarie più contenute. Questo percorso ha caratterizzato lo sviluppo di diverse aziende microelettroniche tra cui quelle europee, che non compaiono oggi tra i primi posti dei produttori globali di componenti microelettronici, essendo questi tipicamente digitali, ma che, nei loro segmenti propongono soluzioni di eccellenza altrettanto rilevanti.

Del resto, nonostante i successi pionieristici di Olivetti nel settore dei primi calcolatori elettronici e l'invenzione, rivelatasi purtroppo pre-

matura del *personal computer* già nel 1965, le tecnologie digitali hanno trovato gli attori principali in Silicon Valley mentre la microelettronica europea si sviluppava a partire da rami di aziende che operavano nel settore dei beni di consumo come Philips (Paesi Bassi) o Siemens (Germania). In Italia Olivetti e Floriani (Telettra) costituiscono nel 1957 la SGS - Società Generale Semiconduttori, in compartecipazione con Farchild, l'azienda americana proprietaria del *know-how* per la realizzazione dei circuiti integrati, pensando soprattutto ad applicazioni nel settore delle telecomunicazioni. Concentrandosi inizialmente sui mercati nazionali di elettronica di consumo (TV, video e telefoni) e sui componenti per le telecomunicazioni e le applicazioni industriali, tutte queste aziende non hanno perseguito la strada della miniaturizzazione e del digitale avanzato ma l'approccio alternativo della diversificazione. Anzi, nel corso degli anni, pur avendo sviluppato prodotti significativi nel settore delle memorie a stato solido (Flash e DRAM), hanno abbandonato questi segmenti per limitare l'esposizione finanziaria necessaria a sostenere la corsa alla miniaturizzazione spinta ed evitare i rischi connessi alle crisi cicliche da sovrapproduzione a cui questi mercati aggressivi sono stati e sono periodicamente esposti.

4. LE FILIERE MONDIALI E IL POSIZIONAMENTO EUROPEO

Oggi i prodotti elettronici sono elementi insostituibili in tutte le catene produttive e il contenuto di semiconduttori nei sistemi continua costantemente ad aumentare. Tuttavia, la produzione elettronica coinvolge decine di attori interconnessi, distribuiti su una scala globale, e ciò rende queste catene di approvvigionamento molto vulnerabili sia a eventi eccezionali come la pandemia sia a tensioni geopolitiche.

Si consideri per esempio un'automobile di media gamma che oggi incorpora circa 50 *chip*, inseriti in varie schede elettroniche. Il produttore di una di queste schede, magari europeo, deve ricevere i componenti da fornitori sparsi sul globo (Giappone, Taiwan, Corea del Sud, US, Europa). Per realizzare ciascuno di questi *chip*, con le sue specifiche caratteristiche, è necessaria una sequenza di contributi manifatturieri e intellettuali riassunti in Fig. 9. E' necessaria una progettazione dedicata (*progettazione chip*) che sfrutta strumenti CAD molto sofisticati. Nella struttura del *chip* possono essere impiegate innovazioni di terze parti (*Intellectual Properties* - IP) come la struttura di un proces-

sore o di una memoria. La fabbricazione dei *chip* su *wafer* di silicio è realizzata in impianti di produzione (*fabbricazione chip*) posti magari a Taiwan o in Corea, costituiti da impianti e macchine specifiche acquisibili da fornitori specializzati (*macchine e strumentazione*). Gli stessi materiali e prodotti chimici impiegati nel processo produttivo sono a loro volta prodotti da aziende specializzate che operano in regime di sostanziale oligopolio (*materiali e wafer*). Al termine della produzione, questi *chip* devono essere provati e assemblati (*test, assemblaggio*), magari in Cina, prima di essere consegnati al produttore della scheda elettronica automobilistica. E' quindi abbastanza comune che ognuno di questi *chip* debba attraversare una decina di frontiere nazionali prima di giungere a destinazione. La sua mancanza o il ritardo nella sua consegna può determinare l'interruzione della produzione di auto, come avvenuto nei mesi di *chip shortage* seguiti alla fine dell'emergenza sanitaria. Analogo impatto sulla economia e sulla sicurezza delle nazioni può determinarsi a seguito della interruzione di forniture causate da tensioni internazionali; ecco quindi che la ripresa della competizione geopolitica tra US e Cina e la guerra Russia-Ucraina ha richiamato l'attenzione dei governi sulla vulnerabilità del settore e ci si è interrogati sul grado di autonomia produttiva delle nazioni e dei sub-continenti.

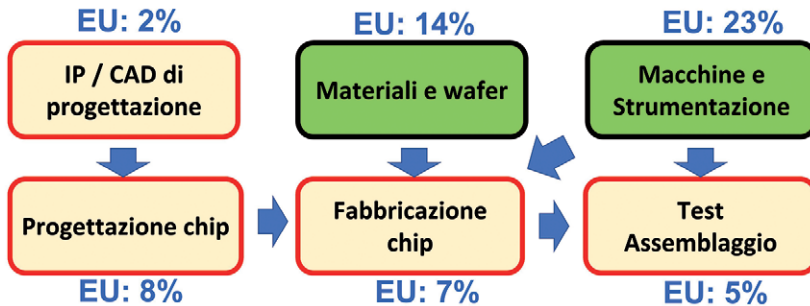


Fig. 9. Segmenti coinvolti nella realizzazione di *chip* microelettronici e percentuale di fatturato delle aziende europee. Figura adattata da [16].

Si è dovuto constatare come nessuna potenza sia oggi in grado di controllare pienamente queste filiere, i cui attori hanno un alto grado di interdipendenza. La Fig. 9 riporta l'incidenza percentuale della produzione europea nei vari segmenti della filiera microelettronica. Per limitarsi alle percentuali, che tuttavia richiederebbero degli approfondimenti ulteriori, si può considerare che il prodotto interno lordo euro-

peo è circa il 17% del prodotto interno lordo mondiale. Usando questa cifra come metro di paragone si può dire che la presenza europea è molto buona nei settori delle macchine di produzione di semiconduttori (23%), è ragionevole nel comparto dei materiali (14%), mentre è molto debole negli altri segmenti, sia sul lato della generazione di IP e della progettazione che per capacità manifatturiera delle fabbriche di *chip*. Oggi il 75% della capacità produttiva di *chip* è in Asia (Giappone 16% – Corea 20% – Taiwan 21% – Cina 15%) contro il 12% in US e il 6% in EU [17, 18]. Se si guarda poi ai nodi tecnologici la situazione appare ancora più sbilanciata. Le fabbriche per i nodi tecnologici avanzati, a più alta densità di transistori, impiegati per i microprocessori e per i sistemi digitali sono di proprietà TSMC (Taiwan) e Samsung (Corea del Sud). Gli Stati Uniti hanno solo il 35% della loro capacità produttiva su nodi <20nm contro il 58% del Giappone e il 37% della Cina. L'Europa ha ben il 75% della sua capacità produttiva su fabbriche con processi più maturi (>40nm) [19].

Questa situazione è il risultato dei cambiamenti determinati dalla specializzazione produttiva del settore partita già alla fine degli anni '80, dalla traiettoria di diversificazione seguita dalle aziende Europee, e dall'affermarsi, negli ultimi venti anni, della globalizzazione. In questo periodo le aziende occidentali hanno progressivamente spostato la fabbricazione in Asia, sulla spinta di alcuni fattori economici come la possibilità di accedere a significative sovvenzioni statali per ridurre i costi di installazione di fabbriche, l'opportunità di realizzare fabbriche in grado di fornire direttamente il crescente mercato asiatico e, solo marginalmente, per il basso costo del lavoro. Parallelamente il modello produttivo è andato modificandosi. Mentre fino agli anni '80 le imprese microelettroniche tendevano a coprire sia la progettazione dei *chip* che la loro fabbricazione, realizzando impianti di proprietà, verso la fine degli anni '80 hanno incominciato a comparire, prevalentemente in Asia, imprese dedicate soltanto alla produzione (*fab*) che forniscono servizi di fabbricazione ad aziende che curano la sola progettazione (*fabless*). Il binomio *fab/fabless* ha continuato a diffondersi nei decenni successivi, in parallelo al processo di trasferimento in Asia della capacità manifatturiera. D'altro canto, come si è detto, le aziende europee hanno progressivamente abbandonato i filoni digitali e si sono concentrati su tecnologie meno aggressive ma capaci di aprire sbocchi nel settore automobilistico, dell'automazione industriale, della sensoristica, della difesa, della sicurezza e dell'aerospazio.

Tuttavia, la lettura del contesto cambia se si considerano gli Stati Uniti e le nazioni politicamente alleate (Europa, Giappone, Corea del Sud). Nonostante la capacità di fabbricazione in US sia scesa al 12%, essa diventa circa il 75% se si amplia lo sguardo al contesto delle nazioni alleate, includendo Taiwan. Il 65% delle aziende di progettazione (*fabless*) sono negli Stati Uniti. Gli Stati Uniti hanno praticamente il monopolio degli strumenti di progettazione e una larga parte delle IP significative. Stati Uniti, Europa, Giappone, controllano una gran parte del mercato delle macchine di produzione. Ad oggi, esiste un solo produttore mondiale di macchine EUV, l'olandese ASML. Si intuisce come il possesso di questa tecnologia sia strategicamente rilevante per controllare lo sviluppo dell'intero mercato dei componenti a semiconduttore. Aspetti critici restano però la capacità produttiva di fabbriche sui nodi tecnologici avanzati, a causa della localizzazione di diversi stabilimenti a Taiwan, e la disponibilità di materiali specifici, come le terre rare, la cui estrazione è dominata dalla Cina.

5. L'AZIONE DEI GOVERNI E IL *CHIPS ACT*

La situazione è quindi molto complicata e passare da una catena di produzione completamente globalizzata ad ecosistemi su base regionale più resilienti alle tensioni geopolitiche richiederà certamente diversi lustri e investimenti ingentissimi. I governi si stanno muovendo in questa direzione anche se con un diverso grado di risorse e chiarezza di obiettivi.

La Cina, nella sua strada verso l'indipendenza strategica dalle forniture e dalle tecnologie occidentali ha avviato già nel 2014 un poderoso programma di investimento nel settore dei semiconduttori nell'ambito del più ampio programma *Made in China 2025* che ha l'ambizione di far passare la Cina dall'essere la più grande fabbrica del mondo ad alta intensità di manodopera a potenza tecnologica avanzata. Si stima che nel quinquennio 2014-19 il piano abbia generato circa 70mld\$ di investimenti per installare fabbriche di semiconduttori e potenziare le capacità di progettazione [20]. Nel quinquennio 2019-24 non è difficile immaginare numeri analoghi.

La risposta americana al *chip shortage* e alle crescenti tensioni geopolitiche con la Cina è stato un *Chips and Science Act* con cui l'amministrazione americana ha stanziato 280mld\$ per il quinquennio 2022-27. La legge è stata promulgata nel 2022 e si propone di sostenere la ricerca

e l'innovazione nei settori di tecnologia avanzata tra cui i semiconduttori. Sul comparto dei semiconduttori si prevedono 39mld\$ in forma di sussidi per incentivare i costruttori di *chip* a installare fabbriche avanzate in US (*reshoring*), 13mld\$ per formare le professionalità necessarie al settore e crediti di imposta sugli investimenti [21].

In Europa la Commissione Europea ha immaginato un'azione analoga elaborando un *Chips Act* che è ora in corso di approvazione [16]. Il piano intende mobilitare investimenti pubblici e privati per 43mld€ su un quinquennio. Tuttavia, le risorse addizionali messe a disposizione sul bilancio dell'Unione sono molto più limitate (circa 3mld€), mentre 11mld€ dovrebbero essere messi a disposizione dagli Stati membri. L'iniziativa prevede la possibilità di erogare aiuti di Stato per incentivare la localizzazione di fabbriche digitali avanzate e anche di altri impianti oggi non presenti sul suolo europeo (*first of a kind facility*). Queste azioni possono concretamente favorire la costruzione di nuovi impianti e il sostegno all'evoluzione dei produttori europei esistenti, tuttavia l'assenza di un significativo tessuto di imprese di progettazione microelettronica, la ridotta presenza nei settori digitali e del software, la minore rapidità con cui l'innovazione scala dallo stato di idee alla generazione di imprese innovative restano aspetti limitanti la dinamica europea.

6. CONCLUSIONI

La crisi pandemica e le tensioni geopolitiche hanno richiamato l'attenzione dei governi e delle opinioni pubbliche sul settore microelettronico. E' cresciuta la consapevolezza della natura strategica di queste produzioni e della necessità di rendere più resilienti alle crisi le catene di approvvigionamento di *chip*. Le azioni messe in campo dai governi delle maggiori economie puntano ad incrementare le capacità manifatturiere e a presidiare le tecnologie abilitanti che permettono la produzione dei *chip* più avanzati. L'Europa da questo punto di vista deve recuperare posizioni e, nonostante le intenzioni, continua a muoversi molto lentamente rispetto alle dinamiche mondiali. Maggiori fondi a disposizione del bilancio comunitario, elaborazione di politiche coordinate, sviluppo dell'ecosistema di progettazione e sostegno all'innovazione restano punti chiave su cui la politica europea deve fare rapidamente passi in avanti per consentire al nostro continente di svolgere un ruolo significativo su questo come su altri temi dell'agenda globale.

RIFERIMENTI BIBLIOGRAFICI

- [1] US 879,532, “*Space Telegraphy*”, US Patent filed January 29, 1907, issued February 18, 1908.
- [2] US 1,900,018, “*Device for controlling electric current*”, US Patent filed March 28, 1928, issued March 7, 1933.
- [3] GB 439,457, “*Improvements in or relating to Electrical Amplifiers and other Control Arrangements and Devices*”, Patent filed March 2, 1934 in Germany, March 4, 1935 in UK, issued Dec. 6, 1935.
- [4] US 2,524,035, “*Three-Electrode Circuit element utilizing semiconductive materials*”, US Patent filed June 17, 1948, issued Oct. 3, 1950.
- [5] US 2,569,347, “*Circuit element utilizing semiconductive materials*”, US Patent filed June 26, 1948, issued Sept. 25, 1951.
- [6] US 3,138,743, “*Miniaturized Electronic Circuits*”, US Patent filed Feb. 6, 1959, issued June 23, 1964
- [7] US 3,025,589, “*Method of manufacturing semiconductor devices*”, US Patent filed May 1, 1959, issued March 20, 1962.
- [8] F. Faggin, and M. Hoff, Jr., “*Standard Parts and Custom Design Merge in Four-Chip Processor Kit*”, *Electronics*, Apr. 24, 1972, pp. 112-116.
- [9] G. Moore, “*Cramming more components onto integrated circuits*,” *Electronics*, vol.38, no. 8, April 19, 1965, pp. pp. 114-117.
- [10] Estratti da “*A Conversation with Gordon Moore: Moore’s Law*”, Intel 2005. http://large.stanford.edu/courses/2012/ph250/lee1/docs/Excepts_A_Conversation_with_Gordon_Moore.pdf.
- [11] K. Rupp, “*42 Years of Microprocessor Trend Data*”, <https://www.karlsruhp.net/2018/02/42-years-of-microprocessor-trend-data/#more-760>
- [12] R. H. Dennard, F. H. Gaensslen, H. -N. Yu, V. L. Rideout, E. Bassous and A. R. LeBlanc, “*Design of ion-implanted MOSFET’s with very small physical dimensions*”, in *IEEE Journal of Solid-State Circuits*, vol. 9, no. 5, pp. 256-268, Oct. 1974, doi: 10.1109/JSSC.1974.1050511.
- [13] International Roadmap for Devices and Systems – 2022 Update. More Moore. <https://irds.ieee.org/editions/2022/more-moore>
- [14] G. Baccarani, M. R. Wordeman and R. H. Dennard, “*Generalized scaling theory and its application to a ¼ micrometer MOSFET design*”, in *IEEE Transactions on Electron Devices*, vol. 31, no. 4, pp. 452-462, April 1984, doi: 10.1109/T-ED.1984.21550.
- [15] “*Multiple Technologies on a Chip, 1985*”. Agrate Brianza, Italy, Dedicated 18 May 2021 - IEEE Italy Section. https://ethw.org/Milestones/List_of_IEEE_Milestones#1980-Present
- [16] European Commission, “*A Chips Act for Europe*”, COM(2022) 45, Brussels, 8/2/2022.
- [17] Jan-Peter Kleinhans, “*The lack of semiconductor manufacturing in Europe*”, April 2021. <https://www.stiftung-nv.de/de/publikation/lack-semiconductor-manufacturing-europe>
- [18] European Semiconductor Industry Association. “*Trends in worldwide semicon-*

- ductor production capacity*”, Brussels June 17, 2012. https://www.eusemiconductors.eu/sites/default/files/ESIA_PR_WWCapacity_2021.pdf
- [19] “*Total Wafer Capacity by Region and Technology Node*”, dati: IC Insights, <https://semiwiki.com/forum/index.php?attachments/monthly-wafer-capacity-2020-icinsights-jpg.350/> in [17].
- [20] “*Global Semiconductors: National Treasures ?*”, Bernstein, Report Dec 2020.
- [21] “*FACT SHEET: CHIPS and Science-Act Will Lower Costs, Create Jobs, Strengthen Supply Chains and Counter China*”, The White House Briefing Room, Statements and Releases, Aug. 9, 2022. <https://www.whitehouse.gov/briefing-room/statements-releases/2022/08/09>