

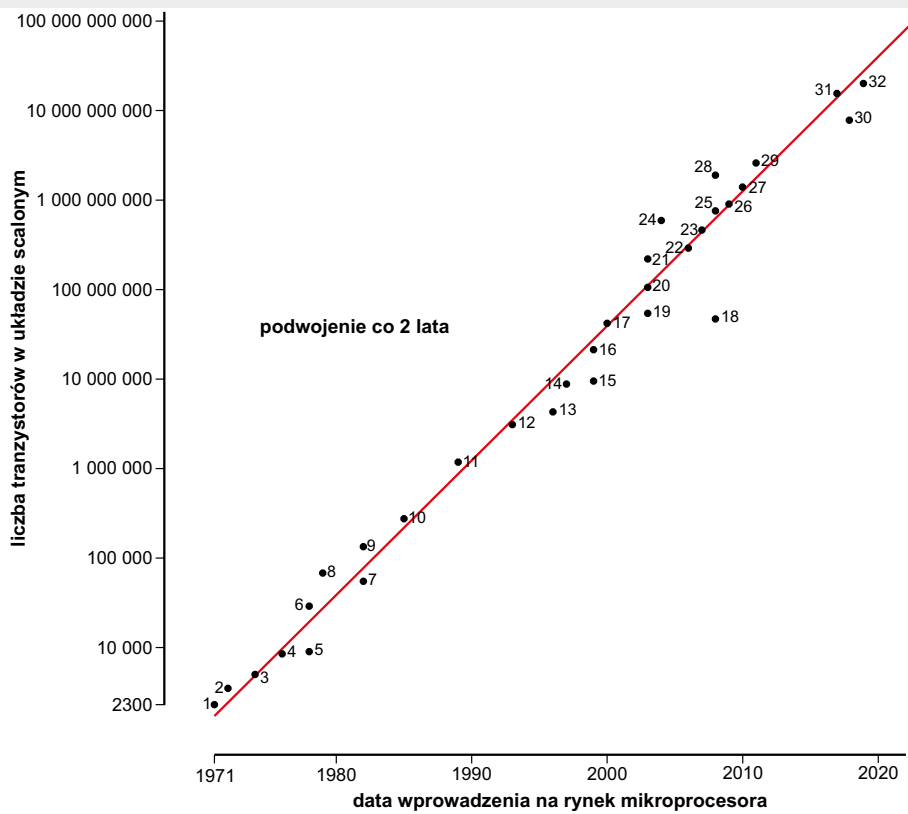
Koniec i co dalej

Czy wyczerpują się możliwości rozwoju mikroelektroniki opartej na technologii krzemowej? A jeśli tak, to co dalej?

Dotychczas przez 60 lat od powstania pierwszych układów scalonych skala integracji, tj. liczba tranzystorów w pojedynczym czipie, podwajała się co 2 lata. Już w połowie lat sześćdziesiątych zauważył tę prawidłowość Gordon Moore, przyszły szef Intelu. Jak widać na **ryśunku 1**, tzw. prawo Moore'a sprawdza się do dziś. Jak długo jeszcze? Oto jest pytanie. Liczba tranzystorów w mikroprocesorach lub pamięciach sięga już dziesiątków miliardów. I co dalej?

Wiadomo, że zwiększanie liczby tranzystorów w układzie scalonym odbywa się dzięki miniaturyzacji tranzystorów. Zatem pytamy, czy miniaturyzacja tranzystorów w układach scalonych zbliża się do swego kresu? Od lat zadawano sobie pytania o fizyczne i technologiczne ograniczenia miniaturyzacji tranzystorów. Wiele prognoz się nie sprawdziło. Kolejne bariery zostały sforsowane. Jednak tym razem to już koniec. Długość kanału tranzystora MOS, szerokości ścieżek i grubości warstw wynoszą już kilka nanometrów, a więc mieści się w nich kilkanaście atomów.

Kolejna redukcja wymiarów o rząd wielkości oznaczałaby, że w długości kanału, szerokości ścieżki lub grubości warstwy mieści się 1 atom, a to nie jest możliwe. Popatrzmy na **ryśunek 2**, jaką drogę przeszła miniaturyzacja przez 50 lat, od roku 1970 do 2019. Wymiar charakterystyczny w pierwszych układach scalonych MOSLSI wynosił 10 μm, by osiągnąć 7 nm w roku 2019. Mówi się, że za rok-dwa możliwa jest jeszcze jedna redukcja rozmiarów do 3 nm, co w Si lub SiO₂ odpowiada dziesięciu



Rysunek 1. Prawo Moore'a ciągle obowiązuje

atomom. I na tym zapewne koniec. Czy to oznacza koniec historii rozwoju mikroelektroniki?

Nie sprawdziły się zapowiedzi końca historii rozwoju społeczno-politycznego (Hegel, Fukuyama) i pewnie błędne byłoby oczekiwanie, że za 2 lata mikroelektronika przestanie się rozwijać. Żeby jednak szukać odpowiedzi na pytanie o przyszłość, trzeba dobrze rozumieć przeszłość. Poświęćmy jej trochę miejsca. Zaczniemy od pytania, dlaczego miniaturyzacja jest utożsamiana z postępem. Po co jest miniaturyzacja? Początek tej historii sięga lat pięćdziesiątych, gdy zaczęły powstawać komputery montowane najpierw z lamp, a potem z tranzystorów. Wtedy pojawiło się pojęcie *tyrania liczb*. Montaż tysięcy tranzystorów był niezwykle pracochłonny i wystarczyło, by jeden z tranzystorów okazał się wadliwy, aby cała praca poszła na marne. Stąd zrodziła się idea, by wytwarzać układ złożony z wielu elementów w jednym procesie technologicznym. W roku 1958 Jack Kilby z Texas Instruments pokazał, że jest to możliwe, demonstrując pierwszy układ scalony wykonany w germanie (był to oscylator składający się z jednego tranzystora ze sprzężeniem zwrotnym RC – **fotografia 3**). Technologia Jacka Kilby'ego nie odegrała żadnej istotnej roli, bo łączył on złotym drutem poszczególne elementy ulokowane na oddzielnych wyspach, co niewiele różniło się od montażu elementów dyskretnych.

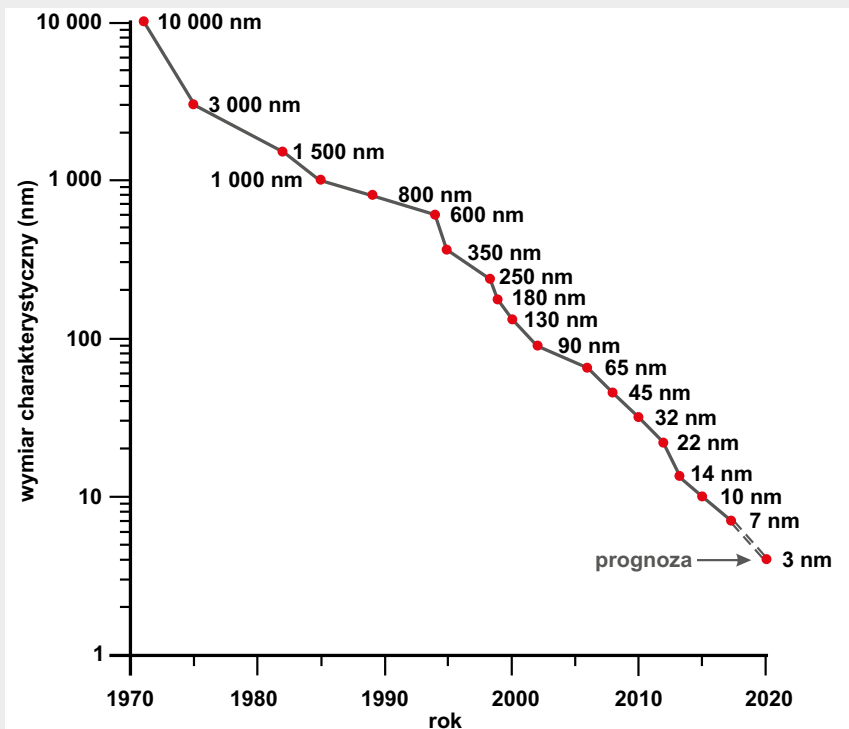
Pół roku po Kilby'ym Robert Noyce w firmie Fairchild zademonstrował układ scalony w technologii planarnej z połączeniami elementów ścieżkami naparowanego aluminium. Był to początek ery technologii planarnej stosowanej do dziś z różnymi modyfikacjami. Był to też początek drogi w redukowaniu kosztów.

Lp.	Nazwa	Liczba tranzystorów	Data wprowadzenia
1	Intel 4004 (4-bit, 16-pin)	2 250	1971
2	Intel 8008 (8-bit, 18-pin)	3 500	1972
3	Intel 8080 (8-bit, 40-pin)	6 000	1974
4	Zilog Z80 (8-bit, 4-bit ALU, 40-pin)	8 500	1976
5	Motorola 6809 (8-bit z niektórymi funkcjami 16-bit, 40-pin)	9 000	1978
6	Intel 8086 (16-bit, 40-pin)	29 000	1978
7	Intel 80186 (16-bit, 68-pin)	55 000	1982
8	Motorola 68000 (16/32-bit, rejestr 32-bit, 16-bit ALU)	68 000	1979
9	Intel 80286 (16-bit, 68-pin)	134 000	1982
10	Intel 80386 (32-bit, 132-pin; bez pamięci podręcznej)	275 000	1985
11	Intel 80486 (32-bit, 4 KB pamięci podręcznej)	1 180 235	1989
12	Pentium (32-bit, 16 KB pamięci podręcznej)	3 100 000	1993

Lp.	Nazwa	Liczba tranzystorów	Data wprowadzenia
13	AMD K5 (32-bit, pamięć podręczna)	4 300 000	1996
14	AMD K6 (32-bit, pamięć podręczna)	8 800 000	1997
15	Pentium III Katmai (32-bit, 128-bit SIMD, pamięć podręczna)	9 500 000	1999
16	AMD K6-III (32-bit, pamięć podręczna)	21 300 000	1999
17	Pentium 4 Willamette (32-bit, duża pamięć podręczna)	42 000 000	2000
18	Atom (32-bit, duża pamięć podręczna)	47 000 000	2008
19	Barton (32-bit, duża pamięć podręczna)	54 300 000	2003
20	AMD K8 (64-bit, duża pamięć podręczna)	105 900 000	2003
21	Itanium 2 McKinley (64-bit, duża pamięć podręczna)	220 000 000	2002
22	2-rdzeniowy Core 2 Duo Conroe (64-bit, duża pamięć podręczna)	291 000 000	2006
23	4-rdzeniowy AMD K10 2M L3 (64-bit, duża pamięć podręczna)	463 000 000	2007

Lp.	Nazwa	Liczba tranzystorów	Data wprowadzenia
24	Itanium 2 (64-bit, 9 MB pamięci podręcznej)	592 000 000	2004
25	4-rdzeniowy Core i7 (64-bit, SIMD, duża pamięć podręczna)	731 000 000	2008
26	6-rdzeniowy Opteron 2400 (64-bit, SIMD, duża pamięć podręczna)	904 000 000	2009
27	6-rdzeniowy Xeon 7400 (64-bit, SIMD, duża pamięć podręczna)	1 900 000 000	2008
28	8-rdzeniowy Xeon Nehalem-EX (64-bit, SIMD, duża pamięć podręczna)	2 300 000 000	2010
29	10-rdzeniowy Xeon Westmere-EX (64-bit, SIMD, duża pamięć podręczna)	2 600 000 000	2011
30	Tegra Xavier SoC (64/32-bit)	9 000 000 000	2018
31	32-rdzeniowy AMD Epyc (64-bit, SIMD, pamięć podręczna)	19 200 000 000	2017
32	AMD Epyc Rome (64-bit, SIMD, pamięć podręczna)	32 000 000 000	2019

Tak jest – podstawowym czynnikiem napędowym miniaturyzacji jest **dążenie do redukcji kosztów**. Pierwszy rynkowy układ scalony wypuszczony przez Texas Instruments w 1960 roku (opracowany przez Jacka Kilby'ego multiwibrator na dwóch tranzystorach, dwóch diodach i ośmiu elementach biernych) kosztował 450 dolarów. Dzisiaj za 1 dolara można kupić układ scalony zawierający ponad miliard tranzystorów. Drugim czynnikiem napędowym miniaturyzacji jest **dążenie do zwiększenia szybkości działania układów**. Im krótszy jest kanał tran-



Rysunek 2. Jak się zmniejszał wymiar charakterystyczny w miarę rozwoju technologii

zystora, tym mniejszy jest czas przelotu nośników od źródła do drenu, a więc większa częstotliwość działania układu. Zatem kierunek i cel rozwoju mikroelektroniki jest oczywisty – coraz mniejsze rozmiary tranzystorów to coraz niższy koszt produkcji liczonego na pojedynczy tranzystor i coraz większa szybkość działania układów scalających coraz większe liczby tranzystorów.

A co nam przeszkadza w realizacji tego celu? Przeszkód jest wiele, a ich działanie ograniczające rozwój ujawniało się z różną siłą na poszczególnych etapach rozwoju. Zmniejszanie rozmiarów od 10 μm do ok. 100 nm, czyli w latach 1970...2000, postępowało dość gładko w wyniku ulepszenia procesów technologicznych, głównie litografii, przy niewielkich zmianach konstrukcji tranzystora poddawanego procedurze przeskalowania (scaling down). W kolejnych latach, przy zmniejszaniu rozmiarów

od 100 nm do 10 nm, trzeba było pokonać kilka ograniczeń fizycznych. Po pierwsze, wzrost szybkości działania układów scalonych zaczęły ograniczać obwody RC tworzone przez pojemności pasożytnicze i rezystancje kontaktów, a więc dalsze zmniejszanie długości kanału nie dawało oczekiwanego przyrostu szybkości działania.

Po drugie, w miarę zwiększania gęstości upakowania tranzystorów i liczby tranzystorów w układzie, rośnie ilość wydzielanego ciepła. Ma to też związek z pojemnościami pasożytniczymi, które zwiększają pobór prądu w stanie dynamicznym, przy przełączaniu inwertera CMOS. Aby zmniejszyć ilość wydzielanego ciepła, trzeba było sukcesywnie zmniejszać napięcie zasilania inwerterów CMOS. Było to też niezbędne dla uniknięcia przebicia skrośnego dren-źródło oraz prądu tunelowego źródło-dren dla zatkanego tranzystora. Dla długości kanału 10 nm stosuje

się napięcie $U_{DD}=0,8\text{ V}$, a dla kanału 5 nm $U_{DD}=0,65\text{ V}$. Tak niskie napięcie U_{DD} wymaga niezwykle powtarzalności i stabilności napięcia progowego U_T . Dla pokonania tych trudności trzeba było wprowadzić istotne modyfikacje w technologii planarnej.

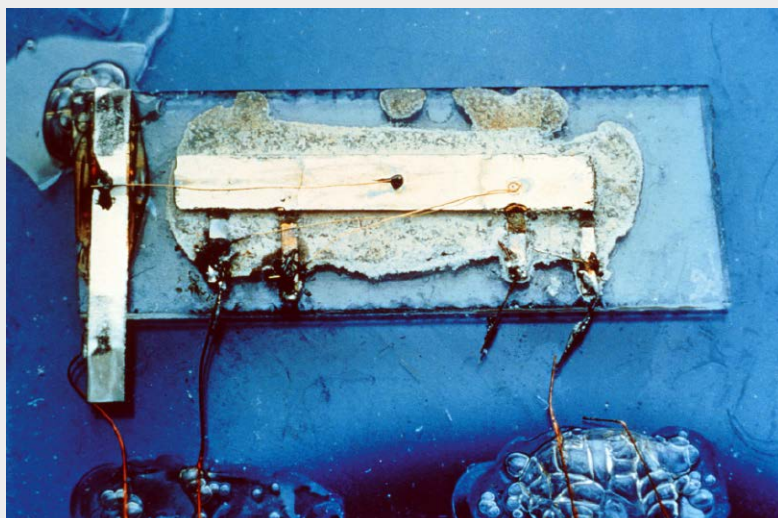
Trzeba było wreszcie sięgnąć po trzeci wymiar. Tak powstała technologia FinFET, w której kanał tranzystora MOS ma kształt pionowej ściany pokrytej branką, która oddziałuje polem z dwóch boków i z góry. Kanał przypomina swoim kształtem płetwę rekina wystającą nad powierzchnię wody, stąd nazwa FinFET. Z tak zmodyfikowaną

technologią planarną udało się „dojechać” do rozmiaru 7 nm i prawdopodobnie uda się jeszcze osiągnąć 3 nm, ale na tym będzie koniec rozwoju mikroelektroniki krzemowej. Czy na pewno? Otóż niezupełnie.

Pojawia się możliwość wytwarzania układów wielowarstwowych. Są już pamięci flash 3DNAND składające się z 96 warstw. Na rok 2021 planowane jest osiągnięcie 256 warstw. Samsung pracuje nad układami 4 Tbit (terabit, czyli 10^{12}) z 410 warstwami.

Przewiduje się, że ta koncepcja pozwoli wytwarzać układy z 1024 warstwami. Jednak wygląda na to, że to łabędzi śpiew technologii krzemowej. I co dalej? Od dość dawna pewne nadzieje wiąże się z wprowadzeniem do gry nowych (starych) materiałów, tj. Ge i GaAs.

Na horyzoncie widać też nowe idee fizyczne. Eksperymenty z zastąpieniem w układach scalonych połączeń miedzianych fotonową transmisją sygnałów mogą doprowadzić do pod-



Fotografia 3. Pierwszy układ scalony wykonany w germanie

wyższenia szybkości działania układów. Od wielu lat wielkie nadzieje łączy się z grafenem i jeszcze większe z tzw. komputerem kwantowym (quantum computing), czyli z nową elektroniką, nazywaną również spintroniką. Historia uczy, że od nowych idei do praktycznych zastosowań na dużą skalę, o ile wszystko idzie dobrze, trzeba ok. 30 lat. Te nowe idee świetnie się zapowiadają już od ok. 10 lat, a więc pozostało jeszcze ok. 20 lat. Przez najbliższe 20 lat będziemy więc mieli ciągłe elektronikę krzemową opartą na transporcie ładunków. Czy będzie to okres stagnacji, gdy prawo Moore'a nie będzie już działać?

Aż nie chce się wierzyć.

Niestaw Nowak