

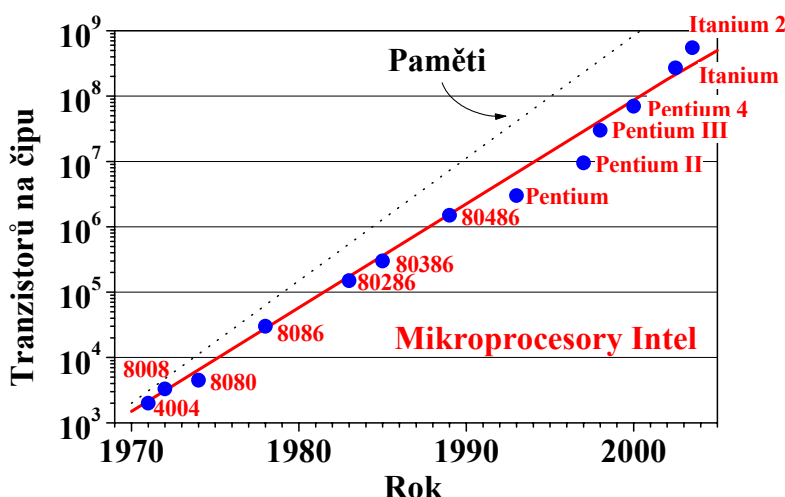
Nové trendy polovodičových součástek - pohledem začátku roku 2005

Jan Vobecký

Katedra mikroelektroniky, FEL - ČVUT v Praze
Technická 2, 166 27 Praha 6, vobecky@fel.cvut.cz

1. Úvod

Letos je tomu čtyřicet let, co Gordon Moore (Fairchild 1957, Intel 1968) prezentoval na základě svého několikaletého pozorování rostoucího počtu tranzistorů na čipu integrovaného obvodu předpověď dalšího vývoje slovy: „Počet tranzistorů v integrovaném obvodu se bude zdvojnásobovat v pravidelných intervalech a tak by to mohlo pokračovat až kam oko dohlédne“. Tento výrok platí s intervalem upraveným na 18 měsíců dodnes a nazývá se Mooreův zákon. V elektronice se uplatňuje v číslicových integrovaných obvodech (IO), jako jsou polovodičové paměti a mikroprocesory, jejichž výkon je odrazem hustoty integrace (Obr.1).



Obr.1: Znázornění Mooreova zákona s využitím mikroprocesorů společnosti Intel.

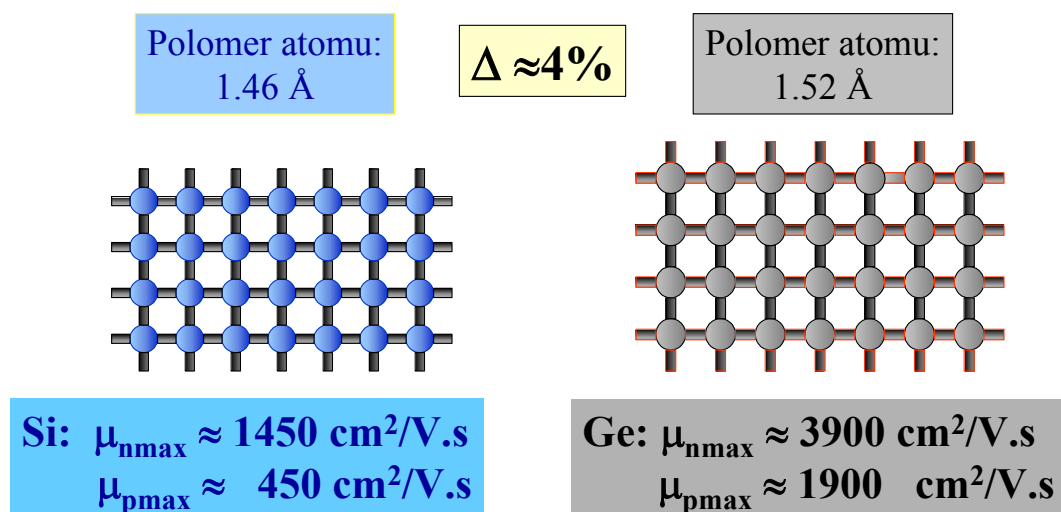
Většina osobností světové vědy a průmyslu se shoduje na tom, že nejdůležitějšími technologiemi uplynulých čtyřiceti let jsou tranzistor (resp. integrovaný obvod), počítač a internet [1]. Za největší překvapení pak považují rychlost vývoje těchto technologií, která odpovídá Mooreovu zákonu, a to prakticky ve všech odvětvích souvisejících s vývojem moderních technologií. Protože tempo růstu dané Mooreovým zákonem je exponenciální funkce, je i pro odborníky zarážející, že se toto tempo dokázalo udržet celých 40 let. Vývoj nových technologií IO ukazuje, že se udrží minimálně po dobu dalších 10 let.

Dosavadní vývoj tedy ukazuje, že vše podstatné se točí kolem tranzistoru. Chceme-li probrat současné trendy vývoje polovodičových součástek, je pak tranzistor logickým předmětem další diskuse. Zvyšování výkonu integrovaných obvodů je možné dosáhnout nejen zmenšováním rozměrů tranzistoru podle obr. 1, ale také volbou materiálů s výhodnějšími

parametry pro dané podmínky. Prvně jmenovanému trendu je v literatuře tradičně věnována poměrně velká pozornost. My se dále zaměříme na druhý aspekt vývoje tranzistoru, který nabývá na stále větším významu zvláště v poslední době, kdy se s minimálními rozměry tranzistoru nezdá se blížíme k fyzikálnímu limitu danému rozměry atomů.

2. Křemík nebo germanium?

Přestože princip p-n přechodu objevil Russel Ohl v *Bell Telephone Laboratories* již v roce 1940 při studiu „zvláštních elektro-optických vlastností“ křemíku, následující objevy tranzistoru s bodovými kontakty (*Point-Contact Transistor*, Bardeen a Brattain, 1947) a tranzistoru s p-n přechody (*Junction Transistor*, Shockley, 1948) byly realizovány na germaniu. Důvodů bylo hned několik najednou. Germanium má teplotu tavení 938 °C a málo reaguje s okolím, takže čištění základního materiálu bylo snadnější. Křemík má teplotu tavení 1415 °C a je chemicky aktivnější, což vedlo k jeho snadné kontaminaci již při výrobě výchozího monokrystalu. Počáteční příklon ke germaniu podpořilo i to, že pohyblivost elektronů a děr v germaniu ($\mu_{n\max} = 3900 \text{ cm}^2/\text{V.s}$, $\mu_{p\max} = 1900 \text{ cm}^2/\text{V.s}$) je asi třikrát větší než v křemíku ($\mu_{n\max} = 1450 \text{ cm}^2/\text{V.s}$, $\mu_{p\max} = 450 \text{ cm}^2/\text{V.s}$). To je dáno tím, že poloměr atomu germania je zhruba o 4% větší než v případě křemíku. Proto jsou atomy germania o trochu dále od sebe než v křemíku a pohyblivost elektronů a děr je několikrát větší, jak ukazuje srovnání na obr.2. S tím souvisí i to, že první tranzistory na germaniu vykazovaly několikrát vyšší mezní kmitočet než křemíkové. Protože byly současně mnohem levnější a spolehlivější, bylo jejich počáteční využití výrazně větší.



Obr.2: Porovnání pohyblivosti elektronů a děr křemíku a germania za ideálních podmínek (nízká dotace příměsí, nízká intenzita elektrického pole) při pokojové teplotě. Nakreslené mřížky nejsou ve správném měřítku.

Zásadním problémem germania však byl a je omezený rozsah pracovních teplot, a to do 70 °C z důvodu malé šířky zakázaného pásu ($W_g = 0.67 \text{ eV}$ při $T = 300\text{K}$), a s tím spojený velký nárůst svodového (zbytkového) proudu při zvýšené teplotě. Tato nespolehlivost byla pro důležité zákazníky, tvořené zejména americkou armádou, do budoucna nepřijatelná. Křemík má oproti germaniu šířku zakázaného pásu téměř dvojnásobnou ($W_g = 1.12 \text{ eV}$ při $T = 300\text{K}$). A protože hodnota svodového proudu je na teplotě závislá exponenciálně, pracují křemíkové tranzistory běžně při teplotách 125 °C i vyšších. Tento zásadní rozdíl poprvé veřejně demonstroval Gordon Teal ze společnosti Texas Instruments v květnu roku 1954 tak,

že před překvapeným publikem střídavě ponořil germaniový a křemíkový tranzistor dvou zesilovačů za jejich provozu do horkého oleje [2]. Zatímco zesilovač s germaniovým tranzistorem přestal okamžitě fungovat, křemíkový pracoval bez problému. Krátce nato firma Texas Instruments uvedla na trh několik typů křemíkového tranzistoru a získala tak na konkurenci náskok několika let. Sluší se poznamenat, že k tomuto zásadnímu výsledku přispěl mimo jiné objev difúze příměsí při vysokých teplotách pracovníky *Bell Labs* v roce 1953 a výroba křemíku firmou Du Pont, později podstatně zdokonalená samotnými výrobci polovodičů. To ale neznamena, že by křemíkový tranzistor okamžitě ovládl trh. V roce 1955 se prodalo ještě 4x tolik germaniových a křemíkových diod, využívaných zejména v diodových číslicových obvodech a usměrňovačích, než tranzistorů [3]. Dominantním materiálem bylo stále germanium, a to i v roce 1959, kdy se tento poměr již téměř vyrovnal. Tento pomalý přechod byl dán tím, že křemík zpočátku neposkytoval srovnatelnou spolehlivost a výtěžnost v masové výrobě jako germanium a křemíkový tranzistor byl zhruba o řád dražší. Co do spínací rychlosti dohnal křemíkový tranzistor germaniový až v roce 1960, a to díky zmenšení jeho rozměrů. To již u germaniového tranzistoru příliš nešlo, a to z důvodu třikrát menší tepelné vodivosti a odtud vyplývajících problémů s odvodem tepla.

Za první integrovaný obvod je považován RC oscilátor Jacka Kilbyho z firmy Texas Instruments (1958) s jedním tranzistorem a několika rezistory a kapacitory vzájemně odizolovanými pomocí MESA struktur a propojenými zlatými drátky vzduchem mimo vlastní destičku. Koncept umístění několika součástek na jednu destičku je dodnes považován za formulaci myšlenky integrovaného obvodu. Technologie jeho provedení však byla pro hromadnou výrobu nepraktická. Integrované obvody se začaly komerčně prosazovat až po realizaci technologie planárního tranzistoru (J. Hoerni, Fairchild, 1958), který inspiroval vývoj technologie planárního integrovaného obvodu (R. Noyce, Fairchild, 1959). Zásadními technologickými změnami byla náhrada MESA struktur a propojovacích drátků difundovanými vrstvami propojenými po povrchu planárními hliníkovými vrstvami s povrchovou izolační a ochrannou vrstvou SiO_2 . To vše se již odehrávalo na křemíku, který se stal v průběhu 60.let hlavním polovodičovým materiálem pro výrobu integrovaných obvodů. Jejich hlavním stavebním prvkem byl přitom bipolární tranzistor podle původního návrhu W. Shockleyho (*Bipolar Junction Transistor* - BJT).

3. BJT nebo MOSFET?

Zatímco na začátku 60. let tvořily IO asi 1% všech používaných součástek v USA, o několik let později to bylo již desetkrát více. Základním stavebním prvkem byl bipolární tranzistor (BJT), na tehdejší poměry dost složitá a záludná součástka s malou výtěžností ve výrobě. Naproti tomu princip polem řízeného tranzistoru (*Field Effect Transistor*), který nastínil Julius Edgar Lilienfeld již v roce 1925, byl v podstatě jednoduchý. Akorát ho nikdo neuměl vyrobit. Až v roce 1960 se to podařilo v laboratořích *Bell Labs* skupině M. Atally, která se po několikaletém úsilí o stabilizaci povrchu křemíku termickým oxidem propracovala nejen k povrchové ochraně proti kontaminaci (a mezitím inspirovala J. Hoerniho k vypracování planárního tranzistoru – viz výše), ale dokázala také o několik řádů snížit hustotu povrchových stavů výchozího křemíkového substrátu. Vývoj MOSFETu (*Metal-Oxide-Semiconductor FET*) do prakticky použitelné úrovně pak po několikaletém úsilí zvládly společnosti Fairchild a RCA. Firma RCA realizovala první řadu CMOS IO (4000), ale do budoucna se poněkud zatlumila realizací státní zakázky CMOS obvodů na safírovém substrátu. Firma Fairchild zase nebrala technologii MOSFETu až tak vážně, i když v roce 1970 vyrobila paměť o kapacitě 1 kbit. Možná i proto v roce 1968 několik pracovníků firmu

Fairchild opustilo a založilo firmu Intel. Ta uvedla na trh v roce 1970 statickou paměť RAM (*Random Access Memory*) s tranzistory NMOS o kapacitě 1kbit a rovněž dynamickou paměť RAM. Tento moment je považován za bod zlomu, protože ukázal ekonomické výhody vyšší hustoty integrace v technologii MOSFET. V rámci vývoje víceúčelového integrovaného obvodu pro kalkulačky mezitím došlo k objevu mikroprocesoru, který se tak přidal k pamětem a nasedl na vlnu Mooreova zákona podle obr.1. Zatímco do roku 1970 dominoval integrovaným obvodům bipolární tranzistor, od tohoto roku již dochází k rostoucí převaze technologie MOSFET.

4. Křemík a germanium?

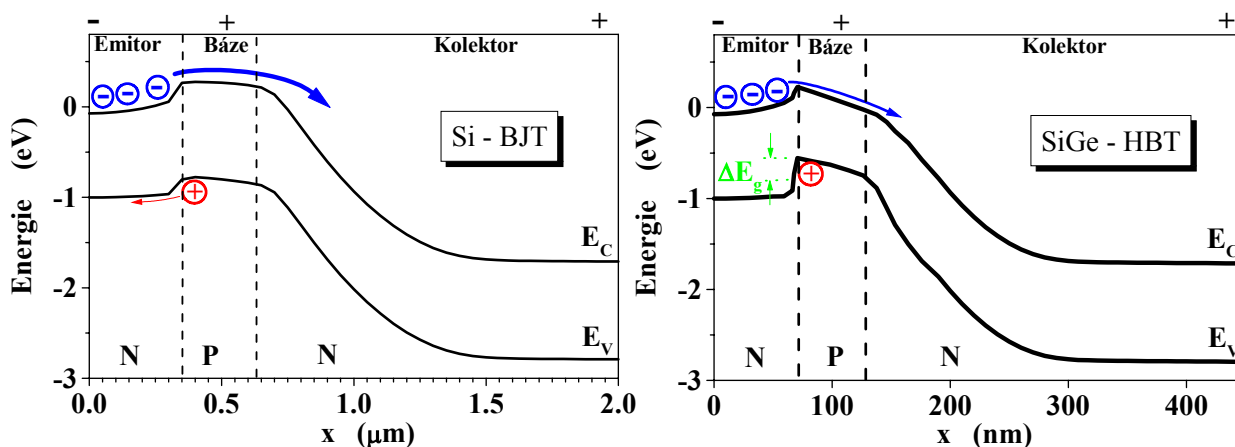
I v posledních letech si křemík udržuje své výsadní postavení, a to i přesto, že mezitím došlo k významnému rozvoji jeho „rychlejších“ konkurentů, jako je arsenid galia (GaAs), indium fosfid (InP) a další. Za hlavní důvod je obvykle udávána nižší výrobní cena hromadné výroby křemíkových čipů, daná možnostmi relativně levné a reprodukovatelné přípravy tenkých vrstev oxidu křemičitého na povrchu křemíku (maskování, hradlový oxid, izolační oxid, pasivace). Samozřejmě, že existují oblasti, jako je optoelektronika, v nichž se jednoznačně prosazují konkurenti křemíku. Za takovou oblast by mohla být považována i komunikační technika v oblasti radiových frekvencí, reprezentovaná např. mobilními telefony, navigačními systémy, ale také internetem, a to z důvodu permanentního tlaku na posun mezních frekvencí z jednotek GHz na desítky a v poslední době až stovky GHz. V tomto ohledu se hranice 200 GHz zdála být pro křemík již nepřekonatelná. V roce 2001 ale firma IBM prezentovala tranzistor HBT (*Heterojunction Bipolar Transistor*) s mezní frekvencí 210 GHz vyrobený v technologii SiGe (křemík - germanium). O rok později pak firma IBM informovala o dosažení mezní frekvence 350 GHz a kompatibility s technologií CMOS (*Complementary MOS*).

Princip bipolárního tranzistoru s emitorem z polovodiče s širokým zakázaným pásem oproti bázi byl patentován W. Shockleym již v roce 1948. H. Kroemer, nositel Nobelovy ceny za přínos v oblasti heteropřechodů (2000), prezentoval v roce 1957 teoretický princip HBT a jeho přednosti. Navrhnul také gradování koncentrace germania v bázi, které se používá pro zavedení přídatného elektrického pole ($\approx 50 \text{ kV/cm}$). To urychluje průlet nositelů náboje bází a zdvojnásobuje tak mezní frekvenci (gradování báze na obr.3 vpravo způsobuje pokles energetických pásů v bázi ve směru x, který prolétající elektron vidí jako narůstající potenciál). Až do sedmdesátých let však neexistovaly vhodné technologie pro výrobu HBT. První HBT tak byl připraven E. Kasperem a jeho kolegy metodou epitaxního růstu (AEG, Daimler-Chrysler) až v roce 1975. V současné době se pro tento účel používá nízkoteplotní epitaxní růst s volitelným profilem koncentrace příměsí a germania v ultratenké bázi tranzistoru.

Na obr.3 vlevo je energetický pásový diagram klasického křemíkového bipolárního tranzistoru NPN (Si-BJT). U něj přikládáme mezi bázi vodivosti P a emitor vodivosti N kladné napětí U_{BE} , abychom snížili potenciálovou bariéru na tomto p-n přechodu a došlo k injekci elektronů do báze. Báze je natolik úzká s ohledem na její dotaci, aby přes ní přešlo co nejvíce elektronů k přechodu báze-kolektor, který je závěrně polarizovaný, a proto odsává elektrony do kolektoru připojeného na kladné napájecí napětí. Protože elektrony injektované do báze rekombinují s děrami, jsou díry do báze doplňovány prostřednictvím bazového proudu. Část děr je bohužel také injektována z báze do emitoru, což je nežádoucí jev, protože snižuje proudový zesilovací činitel. Aby převažovala injekce elektronů z emitoru do báze nad

injekcí děr z báze do emitoru, je dotace báze oproti emitoru zhruba o dva řády nižší. Proudový zesilovací činitel pak nabývá hodnot v řádu stovek. Jeho hodnotu lze zvýšit omezením injekce děr z báze do emitoru.

Na obr.3 vpravo je tranzistor NPN s heteropřechodem (SiGe-HBT). Jeho emitor a kolektor jsou z křemíku, velmi tenká báze (tloušťka v řádu desítek nm) je tvořena epitaxní vrstvou směsi křemíku a germania (SiGe). Injekce děr do emitoru je omezena díky energetické



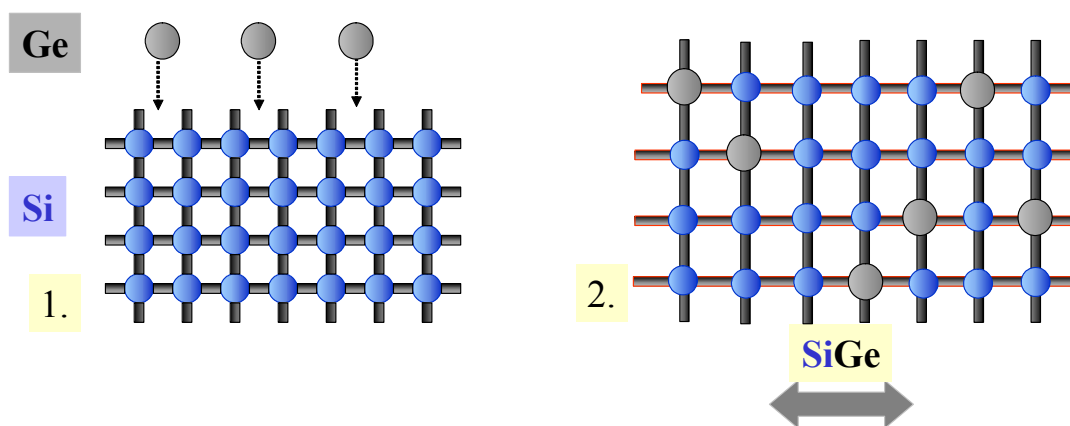
Obr.3: Pásový diagram křemíkového bipolárního tranzistoru (BJT) a SiGe tranzistoru s heteropřechodem provozovaného v normálním aktivním režimu ($U_{BC} < 0$, $U_{BE} > 0$).

bariéry pro díry mezi bází a emitorem (obvykle 0.1 až 0.2 eV). Ta je na obr.3 vpravo naznačena dírou, která není tuto bariéru schopna překonat (energie díry roste směrem dolů, energie elektronu, jehož náboj má opačné znaménko, roste směrem nahoru). Významný nárůst proudového zesílení pak umožňuje použití velmi tenké epitaxní vrstvy báze (typicky 5 – 50 nm) s vysokou koncentrací příměsí (např. bóru). Výsledkem je pak zajištění malého sériového odporu báze, nízké průletové doby báze a vyššího Earlyho napětí (větší odolnosti proti kolísání napájecího napětí). HBT má proto oproti klasickému BJT velmi vysoký mezní kmitočet, nízké šumové číslo a velké zesílení. I když byly HBT s rekordními parametry dosaženy na InP nebo GaAs, zásadní předností SiGe je možnost integrace do výroby křemíkových integrovaných obvodů, a to v kombinaci s technologií CMOS. Výsledkem je technologie SiGe HBT BiCMOS zavedená firmou IBM pro další rozvoj monolitických IO se smíšeným analogovým a číslicovým signálem (1996). Typickou aplikací je přijímací část GPS (*Global Positioning System*) s plně digitální architekturou oproti dřívější analogové verzi [4].

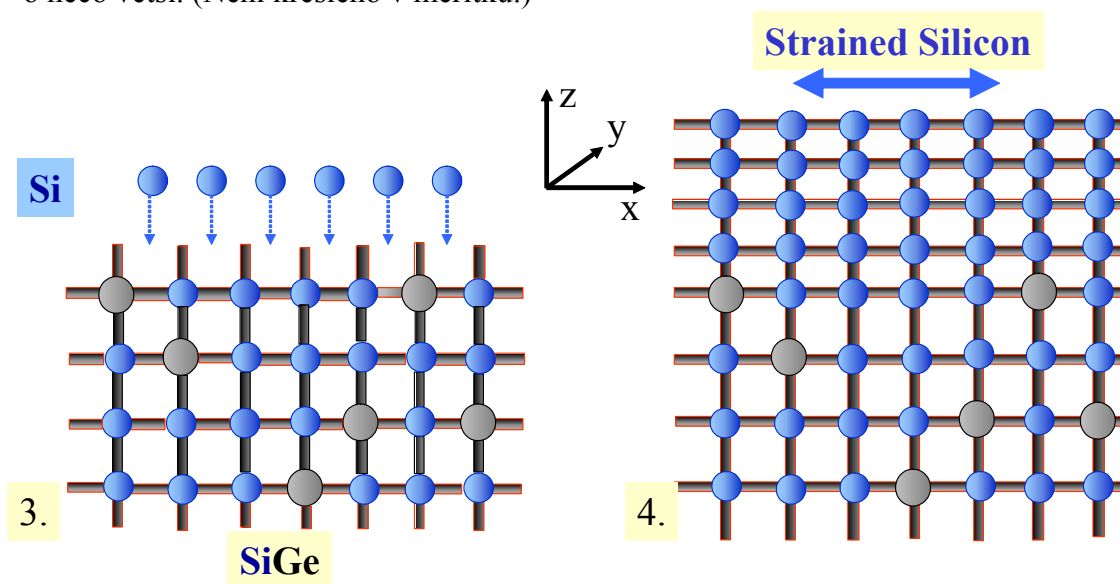
Poznatky získané při vývoji výše uvedených technologií se v poslední době uplatňují i v číslicových obvodech s největší hustotou integrace, jako jsou např. mikroprocesory pro osobní počítače a servery. V nich hraje důležitou roli hodnota pohyblivosti elektronů (kanál N) a děr (kanál P), která se s klesajícími rozměry tranzistorů neustále snižuje, a to díky rostoucím hodnotám elektrického pole kolmého na pohyb elektronů kanálem. Zatímco maximální pohyblivost elektronů, dosahovaná v nízko dotovaném křemíku při pokojové teplotě a nízké intenzitě elektrického pole, je přibližně $1450 \text{ cm}^2/\text{V.s}$, v tranzistoru NMOS s délkou kanálu 130 nm je při jeho provozu přibližně již jen $150 \text{ cm}^2/\text{V.s}$. Jsou proto intenzivně hledány způsoby, jak pohyblivost elektronů a děr v kanálech MOSFETů zvýšit a tím zvýšit i mezní kmitočet mikroprocesorů jako takových. Jedním z kandidátů pro budoucí tranzistory

s nanometrovými rozměry kanálů jsou proto tenké vrstvy germania na izolačních nebo křemíkových substrátech. Aplikace tohoto typu výsledků se ale týká poněkud vzdálenější budoucnosti. Přesto jsme již v současnosti svědky prvního využití křemíkových tranzistorů se záměrně zvýšenou pohyblivostí elektronů a děr, a to v komerčních křemíkových mikroprocesorech. Tento princip si proto stručně popíšeme níže.

V případě tranzistorů SiGe HBT popsaných výše, je báze vytvářena přidáním atomů germania (typicky 10 až 15%). Výsledkem je např. mřížka na obr.4 vpravo, v níž jsou atomy více od sebe (rozdíl činí typ. 1%) a pohyblivost elektronů je proto větší. Výsledkem je větší průletová rychlost bází a tedy i větší mezní kmitočet tranzistoru.



Obr.4: Na obr vlevo nahradíme některé atomy Si atomy Ge a dostaneme mřížku na obr. vpravo. Jelikož jsou atomy Ge o něco větší než atomy Si, je vzdálenost atomů v SiGe o něco větší. (Není kresleno v měřítku.)

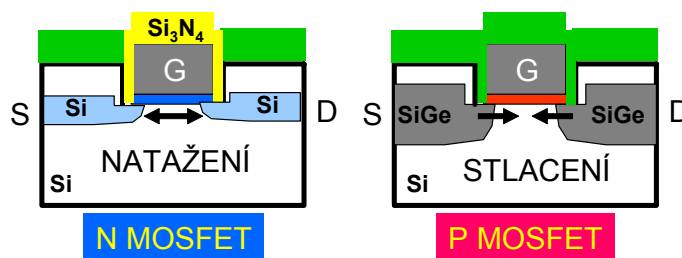


Obr.5: Na vrstvě SiGe (3.) vytvoříme epitaxní vrstvu Si (4.). Protože jsou atomy Si v zákrytu s atomy SiGe, jsou atomy Si v podélném směru o trochu více od sebe, než by tomu bylo v mřížce Si. Dostáváme tzv. „napjatý“ křemík (*Strained Silicon*) s vyšší pohyblivostí elektronů ve směru x a y a nižší ve směru z.

Pokud na mřížce SiGe vytvoříme tenkou epitaxní vrstvu křemíku (obr.5 vpravo), budou atomy křemíku kopírovat polohu atomů substrátu SiGe, a budou proto v podélném směru

(x a y) dále od sebe než by tomu bylo v mřížce křemíku samotného. Na atomy Si tedy bude ve směru x a y vyvíjena síla v tahu, resp. budou v těchto směrech „napínány“. Vzniklá vrstva napjatého křemíku (*Strained Silicon*) pak má ve směru x a y větší pohyblivost elektronů než ve standardním krystalu křemíku. Ve směru z má pohyblivost nižší. Naproti tomu pohyblivost děr ve směru x klesá a roste ve směru y a z. Příčinou těchto změn pohyblivosti μ je změna energetické pásové struktury vrstvy Si vlivem mechanického namáhání, která způsobuje změnu efektivní hmotnosti m_{ef} elektronů nebo děr ($\mu = e \cdot \tau / m_{ef}$). Tento fyzikální princip je znám již poměrně dlouho, a to díky senzorům tlaku.

Zmíněné bi-axiální plošné napínání atomů v destičce Si-SiGe lze s výhodou vyšší pohyblivosti uplatnit i v MOSFETech. Protože hlavní proud vývoje IO spočívá v technologii CMOS, je pozornost věnována především tímto směrem. To je také důvodem nedávného vzniku vývojového zaměření, označeného jako *3D Strain engineering*. Jejím smyslem je co největší současné zvětšení pohyblivosti v kanálu N i P tranzistorů v CMOS technologii, resp. lokální nastavení tlaku a tahu na jednom čipu ve vhodných směrech. Za účelem takového nastavení tažné nebo tlačné síly lze použít vedle zmíněné vrstvy SiGe také tenkých povrchových vrstev Si_3N_4 (krycí izolační vrstva) nebo silicidů, tj. směsí křemíku a kovu vzniklých tepelným zpracováním, tzv. sintrováním. Zatímco na mechanické namáhání vrstev v integrovaných obvodech bylo donedávna nahlíženo jako na čistě nežádoucí jev, v poslední době jsou hledány spíše principy jeho řízeného využití. Uvedené obvody jsou toho důkazem.



Obr. 6: Průřez tranzistorů s namáhaným křemíkem (Strained Silicon) technologie CMOS s minimálním rozměrem 90 nm vyvinuté pro číslicové obvody firmou Intel.

Na obr.6 je princip uniaxiálního napínání a stlačování křemíku používaný firmou Intel v technologii číslicových integrovaných obvodů CMOS s délkou kanálu 90 nm od konce roku 2003 [5]. Oproti dnes již tradiční metodě vytvoření mechanického namáhání pro zvýšení pohyblivosti podle obr.5 je napnutí křemíku v kanálu N provedeno shora krycí vrstvou nitridu Si_3N_4 a stlačení křemíku v kanálu P vrstvou SiGe ze stran. Výhodou tohoto principu je nutnost pouze malé modifikace technologického postupu. Zatímco cenový nárůst výroby je zhruba 2%, nárůst proudu MOSFETem a z něj vyplývající zvýšení maximální spínací frekvence činí údajně až 25%. Tento princip je podle firmy Intel aplikovatelný i na několik příštích generací obvodů CMOS s menším minimálním rozměrem. Dosažení obdobného výsledku bylo prezentováno i společnostmi IBM a AMD koncem roku 2004. S aplikací tohoto principu v procesorech Opteron firmy AMD (90 nm CMOS) bychom se měli údajně setkat již ve druhém čtvrtletí roku 2005.

Závěr

Cílem uvedeného shrnutí vývoje diskretních a integrovaných tranzistorů bylo ukázat, k jakým koncepčním změnám ve využití materiálů může docházet v průběhu vývoje

moderních technologií. Není tomu tak dávno, co jsme studentům na přednáškách vysvětlovali, proč se germanium v polovodičových součástkách využívá již jen u speciálních fotodiod pro měřicí účely a u vysokofrekvenčních hrotových diod. Také si ještě pamatuji, jak nám přednášející na přednáškách o polovodičích koncem sedmdesátých let tvrdili, že „dny křemíku jsou sečteny a že ho zcela určitě již brzy nahradí GaAs.“ Nic takového se však dosud nestalo. Jak totiž praví klasik, *„Je těžké cokoliv předvídat, zejména pak budoucnost.“* Co se týká aplikace nových materiálů, zejména v poslední době je zřejmé, že je třeba dbát velké opatrnosti při předvídání dalšího vývoje. A to se netýká jen materiálů, ale také samotné konstrukce polovodičových struktur. Nezbývá nám tedy než jen pozorně sledovat další vývoj a správně reagovat na probíhající změny. To je ostatně i cílem semináře, na kterém je prezentován tento příspěvek.

Poděkování

Tato práce vznikla s podporou výzkumného záměru MSM6840770017 MŠMT ČR.

Literatura:

- [1] A. Applewhite, The View From the Top, IEEE Spectrum, Vol. 41, No.10, str.16, 2004.
- [2] M. Riordan, The Lost History of the Transistor, IEEE Spectrum, Vol. 41, No. 4, str.36, 2004.
- [3] Facets - New Perspectives On the History of Semiconductors, Edited by A. Goldstein and W. Aspray, IEEE center for the History of Electrical Engineering, Rutgers University, 1997
- [4] B. S. Meyerson, Silicon:germanium-based mixed-signal technology for optimization of wired and wireless telecommunications, IBM J. Res. Develop, Vol. 44, pp. 391 – 407, 2000.
- [5] S. E. Thompson et al, A 90-nm Logic Technology Featuring Strained-Silicon, IEEE Transactions on Electron Devices, Vol. 51, pp. 1790 – 1797, 2004.