

TECHNOLOGICKÉ LIMITY TRANZISTORŮ

Tomáš Podoba

Inženýrská informatika, 2. ročník, prezenční studium
Školitel: prof. Ing. Karel Vlček, CSc.

Fakulta aplikované informatiky, Univerzita Tomáše Bati ve Zlíně
Nad Stráněmi 4511, Zlín 760 05, Česká republika

podoba@fai.utb.cz

Abstrakt. Vývoj čipů, jenž zvládají miliony operací za jednotku času, je přímo úměrný na výrobní technologii tranzistoru. Menší, rychlejší a hustější integrace sebou přináší náročnější procesy, jenž se snaží dodržet Moorova zákona. Základním stavebním prvkem většiny elektrických obvodů je křemík, jehož vlastnosti brzy narazí na technologické limity a pro tranzistory musí být nalezeny nové konstrukční alternativy.

Klíčová slova. *Tranzistor, Limita, High-k, Materiál, Wafer, EUV.*

1 Úvod

Zamyslíme-li se denním bytím člověka žijícího v nerozvojové, civilizované zemi, je na každém rohu, bez jakýchkoliv pochybností, obklopen elektronikou rozmanitého druhu. Dvacáté století bylo považováno za století mikroelektroniky, jenž intenzivně a neodvratně proniklo téměř do všech sfér a činností. Vše způsobila, dnes již pro oko neviditelná, polovodičová součástka zkonstruovaná ze dvou přechodů P a N – tranzistor. Před dvěma lety, konkrétně 16. prosince 2007, oslavil tranzistor díky Bell laboratořím šedesátileté jubileum, přičemž nutno podotknout, že ve svém technologickém principu nedoznal kardinálních změn!

Kvalita, kvantita a ideální propustnost tranzistoru, bude mít vliv na veškeré elektronické součástky dokud lidský tvor neodhalí tajemství kvantové fyziky. Ačkoliv se četné laboratoře předních výrobců integrovaných obvodů snaží imponantně držet exponenciálního trendu vzrůstu počtu tranzistorů na čip (dodržení Moorova zákona [13]), je neodvratně jisté, že se vývoj jednoho dne zastaví díky fyzikálním limitům. V neposlední řadě přijde každá inovovaná či nová generace čipů postavená na rychlejší, respektive menší, technologii na téměř exponenciální vývojové náklady.

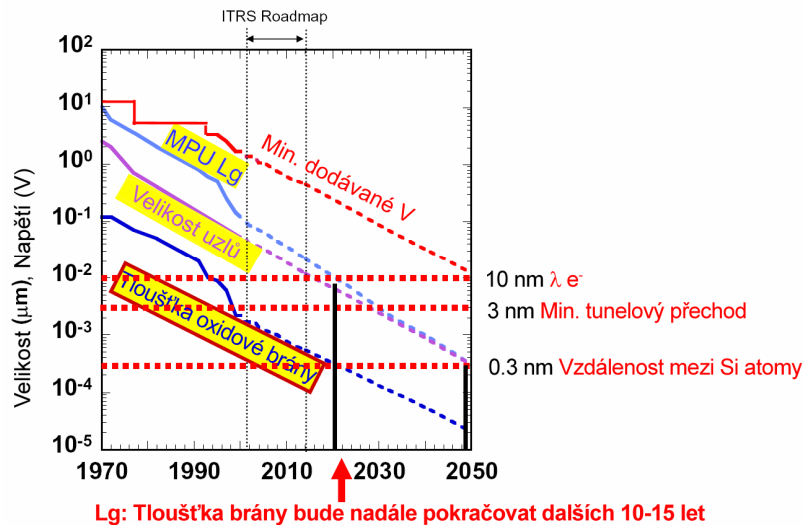
Na výrobním poli se setkáváme s různými technologickými alternativami, jejichž primárním účelem je navyšovat výkonnost mikročipů. Právě vývoj a výzkum je v mnohých případech ovlivněn vratností (zisku) do vložené technologie. Minulý rok byl celosvětovou organizací SEMI [15] zaznamenán 20% pokles.

Prozatímním řešením je výroba tranzistorů založená na materiálech s vysokou hodnotou dielektrika *high-k* pro brány (*gate*) místo současných SiO_2 . Obecně vzato lze říci, že vlastnosti materiálů použitých při konstruování tranzistorů je vertikálou k úspěchu a budou hrát hlavní roli pro další vývoj. Z testovaných kandidátů lze zmínit např. aluminium oxid Al_2O_3 , titanium dioxid TiO_2 , halfium dioxid HfO_2 , zirkonium silikát $ZrSiO_4$, lanthanum oxid La_2O_3 či využití kombinace indium gallium arseniku. Nutno podotknout pořizovací ceny výše zmíněných materiálů jsou v některých případech až desetinásobně vyšší než u SiO_2 tranzistoru. Aktuální 45 nm výrobní technologie používá materiál halfium Hf v konjukci s metalickou bránou. Jenže nalezení vhodného materiálu je pouze

začátek, neboť vhodný materiál je pouze dílčí část komplexního řešení čili nové, lepší výrobní technologie polovodičů.

2 Technologie - materiály

Polovodičová komponenta transistor a jeho princip funkčnosti je více než známý. Transistor si lze představit jako elektronickou hráz rozdělující tok na jedničky a nuly. Čím více transistorů na jednom čipu, tím vyšší výpočetní výkon je dosažen. Dostihovým závodem všech předních výrobců je udržet se trendu navyšování hustoty integrace transistorů, které rovněž způsobují navyšování výkonu. Aktuálně jsou ve výrobě procesory založené na 32 nm technologii, jenž zmenšování bude pokračovat i nadále. Minimální rozměr struktury je roven limitu pro tunelový přechod elektronu (~ 3 nm), přičemž vlnová délka elektronu je 10 nm a vzdálenost atomů 0.3 nm. [1, 10]



Obrázek 1: Počet čipů na wafer.

Z výše zmíněných čísel lze jednoduše odvodit teoretické mezníky pro konstruování transistorů s použitím SiO_2 (obrázek 1). Proto jsou zkoumány různé varianty materiálů, které již nahrazují a nahradí křemík a pošlou tak staré komponenty do často frázovaného „křemíkového nebe“. Materiály mající vysokou hodnotu dielektrika *high-k*, Lanthanoidy (*La, Ce, Yb, Lu...*) či Aktinoidy budou jistě dalšími kandidáty v pořadí. Jisté je jedno, pořizovací cena těchto materiálů na úkor konstrukce eventuelních budoucích transistorů, bude velkou neznámou.

Stále bude ode dneška následovat 4 až 7 (generačních) technologických cyklů v mikroelektronice, než dosáhneme 11 ~ 5 nm výrobní technologii. Vzhledem k tomu, že každá generace trvá v průměru tři roky, lze předpokládat dosažení minimální velikosti transistorů - limitu - z materiálů dostupných na Zemi někdy kolem roku 2025 až 2030. Předpokladem bude taktéž vývoj alternativ jako vertikální a hybridní transistory, 3D logické obvody, nanowire/tube MOSFET (blízká budoucnost) či jiných přenosných kanálů, bio-systémy (obrázek 4) [8, 9].

3 Velikost waferu

Aby byla daná výrobní technologie co nejefektivnější, je nutno během jednoho procesu vyrobit co nejvíce čipů na wafer C_w , kde je plocha waferu dělena plochou samotného čipu (1). [18] Silikonový

respektive jakýkoliv wafer musí být téměř 100% (ideální, ovšem nereálný případ) čistý krystalický materiál s tolerovanou odchylkou 0,01 %.

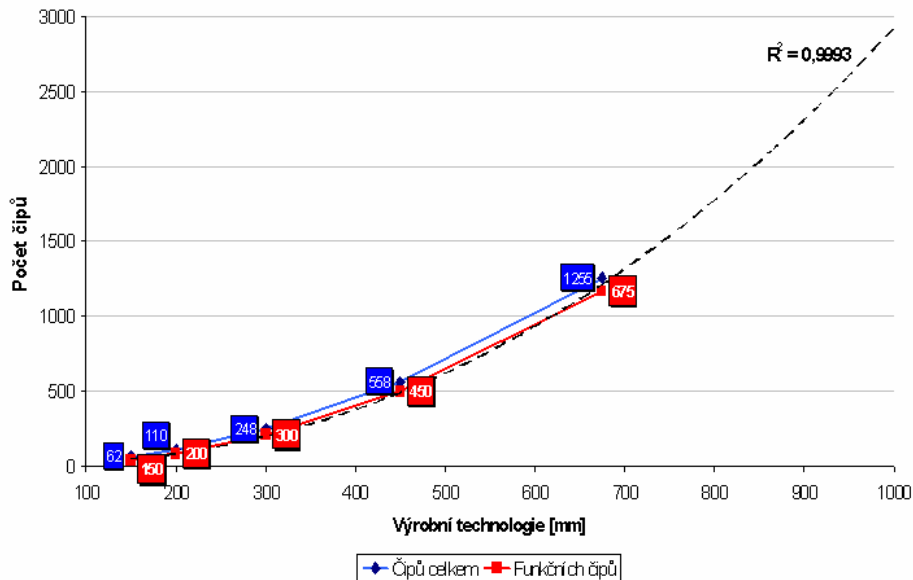
$$C_w = \left[\frac{\pi \times \left(\frac{D_{wafer}}{2} - E_{wafer} \right)^2}{A_{chip}} - \frac{\pi \times D_{wafer}}{\sqrt{2 \times A_{chip}}} - T_{wafer} \right] \quad (1)$$

Kde:

- C_w = počet čipů na wafer
- D_{wafer} = průměr waferu [S, mm²]
- E_{wafer} = kritická okrajová plocha waferu [mm²]
- A_{chip} = velikost čipu [mm²]
- T_{wafer} = další omezující koeficient (okraje, plocha, zmetkovitost...)

Během náročného výrobního procesu čipů, dochází k defektům čipů především v krajních plochách E_{wafer} . Uváděná míra defektu čipů se pohybuje od 0,004 – 0,002 na mm², přičemž efektivní počet funkčních čipů je přímo úměrný velikosti waferu, kde bereme v úvahu precizní výrobní technologii. Právě výrobní preciznost je stěžejním faktorem pro přední polovodičové konsorcia. Náklady na výzkum a vývoj každé nové technologie umožňující výrobu menších čipů na větších waferech jsou rentabilně spekulativní.

Z rovnice (1) byly vypočítány následující hodnoty (obrázek 2), čísla hovoří za vše. V případě 200 mm technologie je v průměru vyrobeno 80 funkčních čipů na wafer což odpovídá 75% efektivnosti. Při 300mm potom 205 čipů ~ 83% efektivnost, 450 mm (aktuálně ve vývoji a předpokládaná sériová výroba během roku 2009) přinese až 500 čipů ~ 89% efektivnost a konečně 675 mm až 100 čipů ~ 92% efektivnost. Kalkulováno bylo se stejnou velikostí čipu 285mm², přičemž počet tranzistorů na čip by se zvyšoval podle Moorova zákona [1].



Obrázek 2: Počet čipů na wafer.

4 EUV

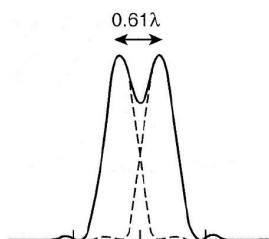
Nebo-li Extreme ultraviolet lithography (EUV nebo EUVL) je vyspělá technologie pro výrobu mikroprocesorů založená na vypalování „obvodového designu“ - předlohy pomocí intenzivních ultrafialových paprsků přímo na silikonový wafer. Hlavní výhodou EUV technologie je možnost nastavování rozdílných vlnových délek s různými parametry, kde dochází ke zmenšování poměru mezi předlohou integrovaného obvodu na masce a výsledným obrazem na waferu [12].

U klasické výrobní technologie polovodičů pomocí optické litografie je stěžením požadovaného výrobního výsledku rozlišení R . Příslušný poměr je zvyšován kombinací redukované vlnové délky paprsku λ (ve vakuu), navyšováním indexu numerické apertury NA projekčního paprsku a technologie *image enhancement* jak je patrné z rovnice (2).

$$R = \frac{k_1 \lambda}{NA} \quad (2)$$

Z rovnice (1) lze vyjádřit fyzikální limity optické litografie přímo závislé na:

- Indexu NA - Teoretické maximum NA je okolo 1.4
- Hodnota k_1 vyplívá z difrakčního limitu paprsku, který je definován vztahem $\lambda / 2$ jenže z praktického hlediska je mnohem přesnější vyjádření rovnicí $k_1 * \lambda / NA$, kde $k_1=0.61$ (obrázek 3)



Obrázek 3: Rayleigh kritérium $k_1 = 0.61$.

Z výše uvedených faktů je patrné, že musela přijít nová technologie – EUV, jenž využívá ultrafialového paprsku s extrémní vlnovou délkou $\lambda=13.5$ nm pro konstruování integrovaných obvodů – čipů. Díky EUV jsme schopni vyrábět čipy s technologií menší jak 30nm (první čipy s 32nm – INTEL a 22nm SRAM od IBM; prozatím na 300mm waferu). [2, 3]

5 Cíle disertační práce

Cílem disertační práce je vytvoření GPU benchmarku založeného na real-time renderingu komplexního uzavřeného meshe vygenerovaného pomocí strange atraktorů. Atraktor musí být vytvořen jako 4D object pomocí quaternion algebry. Vstupem každého testu bude variabilita uživatelských či přednastavených parametrů ovlivňující komplexnost výstupního objektu a tedy celé výpočetní operace. Součástí bude zakomponováno efektivní zpracování efektů jako ray-tracing, anti-aliasing, anisotropní filtrování a další pro hodnocení výkonnosti GPU. Nejnáročnější procedurou bude návrh a realizace algoritmu pro vytváření opravdového uzavřeného třírozměrného objektu,

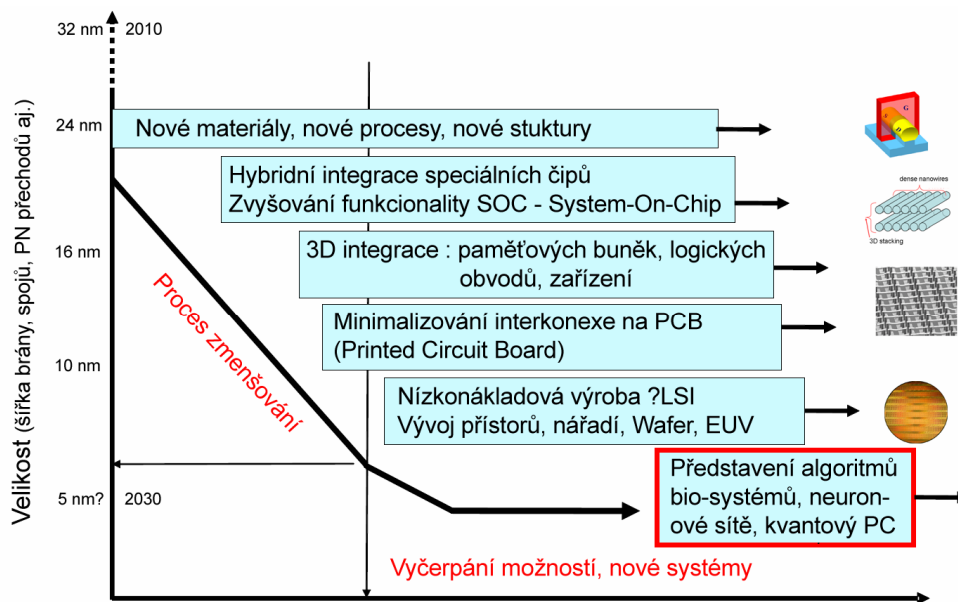
který bude následně hodnocen. Real-time nastavování vstupních parametrů prozatím není smýšleno pro obrovskou náročnost výpočtů.

Dílo technologické limity tranzistorů vzniklo na základě mého dvouletého vyučování předmětu Pokročilé architektury procesorů a taktéž jako zkouška z předmětu návrh a implementace číslicových obvodů v rámci prezenčního doktorského studia na FAI UTB ve Zlíně. Jednotlivé části jsou potom přednášeny v seminářích předmětu nebo samotné přednášky. Autor je interesován v této problematice, bohužel není možné se s ní setkat přímo ve výrobě.

6 Závěr

Výrobní limity tranzistorů nelze přesně určit, přesto jsou známy konkrétní fyzikální mezníky, které v případě nenalezení vhodných alternativ, budou dosaženy. Všeobecně vzato lze rozdělit limity tranzistorů na fyzikální, materiálové, chemické, výrobní a limity struktur. V případě fyzikálních limit je brána v úvahu minimální potřebná tepelná energie pro nosiče elektrického náboje, jenž je limitována na minimální spínací energii pro jedno binární přepnutí na 2 až 4 kT (0,05 až 0,1 eV pro $T=300\text{ K}$, k - Boltzmannova konstanta). Zvyšování pracovní rychlosti (f – frekvence) způsobí vyšší až nepřijatelné parazitní elektrický příkon, což limituje hustotu prvků a tedy celkové integrace obvodu.

Materiálové limity jsou přímo závislé na potřebné spínací energii, kdy např. u high-k materiálů je potřeba několikanásobně nižší spínací energie na úkor ceny. Křemík je 3x levnější než *GaAs*. Materiály jako *Hf*, *La*, *Lu* jsou budoucím stavebním prvkem tranzistorů. Nutno také neopomenout chemické limity v rámci dlouhodobé stability struktur jako difúze aktivních příměsí, elektromigrace atomů.



Obrázek 4: Předpokládaný vývoj velikosti elektronických součástek.

A konečně limity *Si* struktur. Narážíme na limitní tloušťku hradlového oxidu 3 nm při napájecím napětí 1,5 V. Obecně lze říci, že minimální výrobní technologie MOSFET tranzistorů s použitím *Si* je 5 nm (přesněji 5.5 nm). Jenže, všechny nové objevy popř. technologie musí být zkonstruovány a proto je zcela klíčový vývoj zařízení, jenž bude schopno tak malé prvky vyrábět. Využití EUV paprsků, dokonalost masky a nastavování příslušných parametrů (k_1 , NA) povede k dokonalejším, rychlejším a především rentabilnějším produkcím čipů na větších waferech (450 mm ~ 675 mm).

Proto je nutno se zaměřit na vývoj nových alternativních prvků, které budou mít stejnou roli – roli hráze – pro binární přenos v elektrických obvodech. Mezi alternativy patří vývoj vertikálních či 3D logických obvodů, interakci nano-trubiček (nanowire/tube) nebo bio-systému (obrázek 4). A to do té doby, než se lidstvu podaří realizovat plně funkční kvantový počítač.

Reference

- [1] 4. celostátní seminář Perspektivy elektroniky, sborník přednášek, Rožnov pod Radhoštěm, květen 2005
- [2] Benschop, J. P. H., 32 nm: Lithography at a Crossroad, Future Trends in Microelectronics, Edited by Luryi S., Xu J., Zaslavsky A., 2007, ISBN 0-471-48
- [3] Benschop, J. P. H., Limits and Alternatives to Optical Lithography, Future Trends in Microelectronics, Edited by Luryi S., Xu J., Zaslavsky A., 2004, ISBN 0-471-48405-8
- [4] Bevington, P. R., Robinson, D. K.: Data Reduction and Error Analysis for the Physical Sciences, 2002, ISBN 0072472278
- [5] Brillouët, M.: Physical Limits of Silicon CMOS: Real Showstopper or Wrong Problem?, Future Trends in Microelectronics, Edited by Luryi S., Xu J., Zaslavsky A., 2007, ISBN 0-471-48
- [6] Hawrylak, P., Quantum Computation – Future of Microelectronics?, Future Trends in Microelectronics, Edited by Luryi S., Xu J., Zaslavsky A., 2007, ISBN 0-471-48
- [7] Mark T. Bohr, Robert S. Chau, Tahir Ghani, and Kaizad Mistry: IEEE Spectrum: The High-k Solution, dostupné na: <http://www.spectrum.ieee.org/semiconductors/design/the-highk-solution>
- [8] International Technology Roadmap for Semiconductors 2006, dostupné na <http://www.itrs.net/>
- [9] Iwai, H., The Future of CMOS Downscaling, Future Trends in Microelectronics, Edited by Luryi S., Xu J., Zaslavsky A., 2004, ISBN 0-471-48405-8
- [10] Iwai, H.: Downsizing of transistors towards its Limit, NIT Calicut, January 2009.
- [11] M. Passlack, R. Droopad, K. Rajagopalan, J. Abrokwah, and P. Zurcher.: High Mobility III-V MOSFET Technology, 2006
- [12] Melzner H.: Smaller is Better? Maximization of Good Chips per Wafer by Co-Optimization of Yield and Chip Area, Advanced Semiconductor Manufacturing Conference, 2006. ASMC 2006. The 17th Annual SEMI/IEEE, May 2006, ISBN 1-4244-0254-9
- [13] Moore, G.: “Cramming more components onto integrated circuits”, Electronics 38, 114 (1965)
- [14] Nishi, Y.: Scaling Limits of Silicon CMOS and Non-Silicon Opportunities, Edited by Luryi S., Xu J., Zaslavsky A., 2007, ISBN 0-471-48
- [15] Peide, D. Ye.: Beyond Silicon’s Elemental Logic, IEEE Spectrum, dostupné na: <http://www.spectrum.ieee.org/print/6634>
- [16] SEMI is the global industry association serving the manufacturing supply chains for the microelectronic, dostupné na <http://www.semi.org/>
- [17] Solomon, P. M.: Strategies at the End of CMOS Scaling, Future Trends in Microelectronics, Edited by Luryi S., Xu J., Zaslavsky A., 2004, ISBN 0-471-48405-8
- [18] Weems, C.: Technology and Economics of High Performance Systems, Course: CmpSci635A, University of Massachusetts, Dept. of Computer Science