FinFET technology and its advancements- A survey

Manisha Jayson, Anitha Senthil Kumar

Abstract— Planar CMOS transistor scaling presents several difficulties to s ecure an acceptable gate to channel control. Controlling current leakage when the tr ansistors are switched off is important to minimise power consumption in computer and mobile applications .The in-use 20nm process node has introduced a new set of challenges, including double patterning and leaky transistors due to short channel effects. While the planar FET may have reached the end of its scalable lifespan, the semiconductor industry has found an al ternative approach with FinFETs, which allows increasing the gate s caling above the planar transistor limits, sustaining a steep sub threshold slope, better performance with bias voltage scaling and good m atching due to low doping concentration in the channel.

Index terms—scaling, static leakage, power consumption, short channel effects

INTRODUCTION

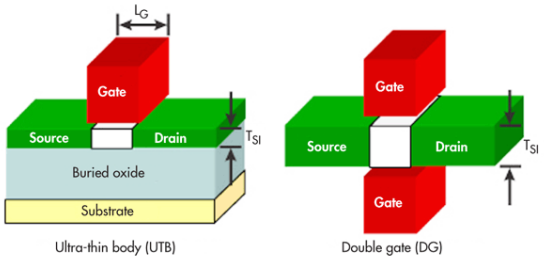
Modern mobile and computing devicesinnovate at a dramatic rate delivering more performance in smaller form factors with higher power efficiencies. Since the inception of the integrated circuit industry, design metrics such as performance, power, area, cost, and time-to-market have remained the same. In fact, Moore's law is all about optimizing those parameters to produce the smallest possible transistor size with each new technology generation.[1]When Gordon Moore came up with his law back in 1965, he perceived a design of about 50 components. Today's chips consist of billions of transistors, and design teams strive for 'better, sooner, cheaper' products.TheFinFET transistor structure promises to rejuvenate the chip industry.[9] It will do so by reducing the short-channel effects[13].

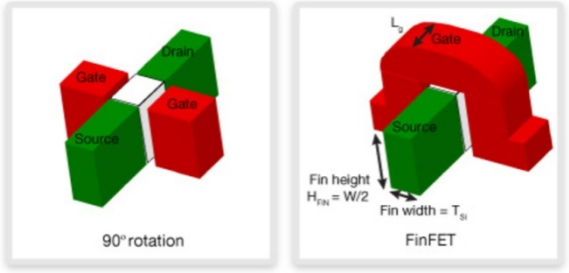
This paper is organized as follows. Section II describes the history of FinFET and its technology. Section III clarifies the design and working mechanism. Section IV explains the growth potential and Section V shows a comparison with non FinFET devices. Section VI looks at the design verification aspects, Section VII looks at the future path of FinFET technology. Finally, Section VIIIexplains the existing challenges for designers.

HISTORY

The research of multi-gate MOSFET happened about quarter century ago, in late 1980s.The first multi gate transistor was that proposed by Hieda et al in 1987. From it, designers realized that fully depleted body of silicon based transistor helps improve switching due to lessened body bias effect. After two years, Hisamoto et al demonstrated an early version of FinFET in bulk silicon, called DELTA.[2] In 2004, a University of California Berkeley team, led by Dr.Chenming Hu, proposed a new structure for the transistor that would reduce leakage current.

The Berkeley team suggested that a thin-body MOSFET structure would control short-channel effects and suppress leakage by keeping the gate capacitance closer to the whole of the channel. The proposed structures were as follows





First FinFET on SOI substrate was published a decade later. SOI also enabled horizontal gate-all-around (GAA) transistor creating a precursor to silicon nanowire

devices. Stacking more than one nanowire on top of each other demonstrated increased drive current capability for a given size of a transistor.

1 DESIGN

1.1 TCMS gate control

FinFET uses a mechanism called threshold voltage control through multiple supply voltages (TCMS).[12] The threshold voltage at each FinFET gate is not only controlled statically through the control of process parameters, such as channel-dopant concentration or the value of the gate work function, but also dynamically through the application of a voltage to the other gate (gate–gate coupling).

Here is a FinFET front gate threshold voltage and the relationship between the gate voltage, the following is a simple approximation of this relationship.

$$V_{th_{gf}} \approx \begin{cases} V^0_{th_{gf}} - \delta(V_{gbs} - V_{th_{gb}}) & \text{if } V_{gbs} < V_{th_{gb}} \\ V^0_{th_{gf}} & \text{otherwise} \end{cases}$$

where s represents the source of the FinFET, and δ is an effective number determined by the ratio of the gate and body capacitance, $V^0_{th_{gf}}$ which is the $V_{th_{gf}}$ minimum, and the above formula is given according to the n-type FinFET device, but by changing the sign can be suitable for p-type FinFET devices. If the two gates of the FinFET device are connected together, then one of the two gate voltages changes the threshold voltage of both gates to change simultaneously with it. As indicated by the above formula, the coupling between the grids is only found in the weak inversion state. In the strongly inverted region, the presence of charge in the inversion layer in the channel blocks the connection between the gates of the FinFET so that no coupling between the gates is observed.[3]

In Figure 1, the inverter a is called the high-Vdd inverter, precisely because it is connected to the operating voltage V^H_{dd} and V^H_{ss}. The inverter b is a low-Vdd inverter because it is connected to the operating voltage V^L_{dd} and V^L_{ss}. V^H_{dd}, V^H_{ss} And V^L_{dd} the values are 1.08V, -0.1V and 1.0V, respectively. VLss refers to ground. Thus, the inverter b here is in overdrive

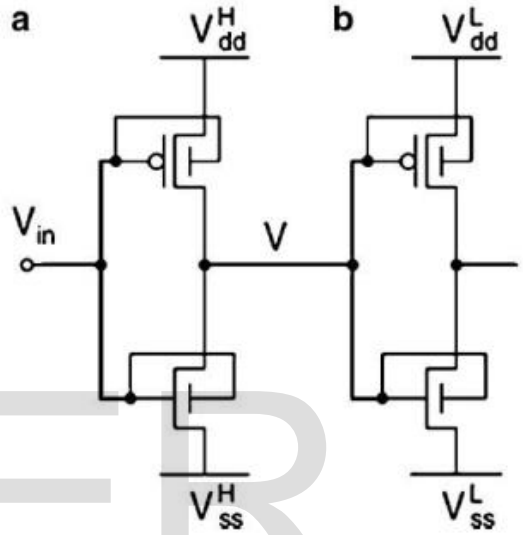


Figure 1 TCMS principle

It can be seen from this point that any V^H_{dd} connected inverter is connected at the to same time V^H_{ss} low.To illustrate how TCMS works, we keep Vin in Figure 1 at a logic zero level, so here is V = V^H_{dd}. Therefore, the sub threshold leakage current in the inverter b is determined by the leakage current of the p-type FinFET therein. However, the delay of the inverter is largely determined by the current flowing through the n-type FinFET.

According to the above formula, the threshold voltage Vth of the p-type FinFET increases due to the presence of a reverse bias voltage of 0.08 V on its gate, thus reducing the sub-threshold leakage current. At the same time, the n-type FinFET device is in a strong reverse state, and no significant change occurs. However, the addition to this inverter is a forward bias voltage of 1.08V, which is higher than the normal case of the

inverter input gate voltage of 1.0V, which leads to a larger drive current. Similarly, a logic 1 level at the input of the circuit can result in a decrease in the leakage current of the n-type FinFET device and an increase in the drive current of the p-type FinFET device.

1.2 Fin shape

The first FinFET based high performance logic product - Intel's 22nm node microprocessor has been built with FinFET sidewalls sloping at about 8 degrees from vertical. Such shape has several practical reasons for manufacturability of this technology[5]Fins with lower aspect ratio (height: width) are more robust mechanically and so less exposed to damage processing. Sloping sidewalls promise better fill of trenches between fins with fin isolation dielectric.They also have a major drawback – poor short channel control near the bottom of the fin. Such fins would usually require additional doping to lessen this problem. Thus causing increased random dopant fluctuation. The drawback of sloping fin sidewalls become serious with scaling gate length and will need a more vertical shape.

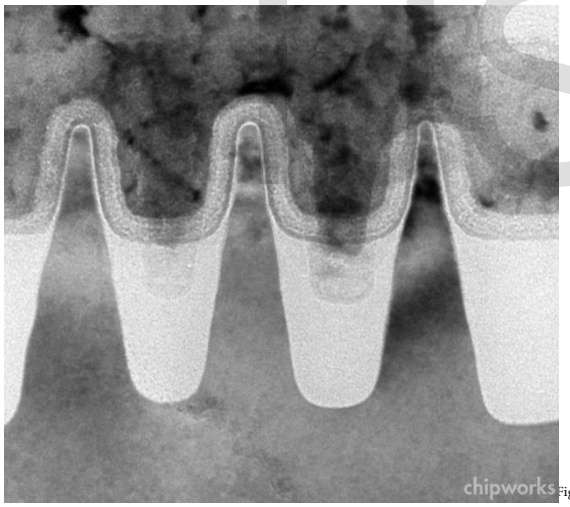


Fig 2. TEM image of an NMOS cross section with fins

1.3 Doping

Some light doping is required to set threshold voltages for better control of leakage current. Those doping are done by implantation. Source/drain doping requires high doses of dopant, thus increasing series resistance. This causes implant damage in the fins .High temperature (300-400C) implants called plasma-based doping or monolayer doping methods deliver dopants with less damage to the fin. Alternatively, doped epitaxial material is deposited in source/drain area to deliver the dopant. It can be done with or without removing the fin in source/drain area prior to epitaxial material.

1.4 Fin patterning

In order to match the effective width of a FinFETdevice, their fins needed to be very tall. Usually, formation of two fins per minimum pitch allows reasonable fin aspect ratio that meets or exceeds effective width of corresponding planar device. Traditional lithographic patterning has been found to be erroneous resulting in fin length variability.

1.5 Parasitic capacitance

FinFET has inherently higher parasitic capacitance than planar device. It consists of gate-to fin capacitance between part of the gate above the fin and the top of the fin. Parasitic capacitance decreases with decreasing fin pitch and increasing fin height, per unit effective device width .Bulk FINfet junction capacitance between source/drain area and device well/substrate could be several times smaller than in planar devices.

1.6 Design Tools

[4]Most of Electronic Design Automation (EDA) tools needs to be adapted for FinFET designs. This process has been largely completed and tools are available from key vendors like Synopsis, Mentor Graphic and Cadence.

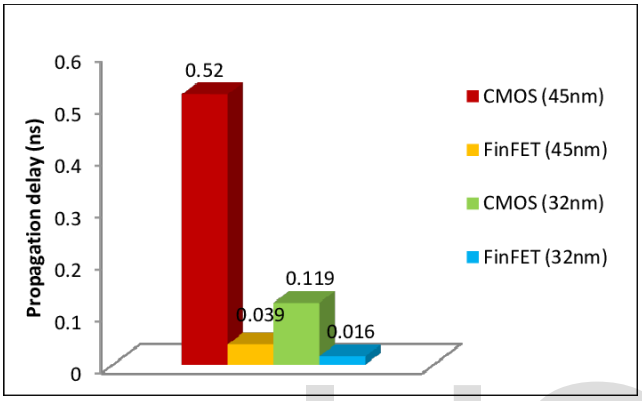
2 GROWTH POTENTIAL

The FinFET technology market is expected to grow from USD 4.91 Billion in 2015 to USD 35.12 Billion by 2022.Although the FinFET technology market is currently dominated with laptops and tablets end user segment, the smartphones and wearables segments are expected to gain traction and grow at the highest rate during the forecast period. The key players in the FinFET technology market include Intel (U.S.), TSMC, Ltd. (Taiwan), Samsung (South Korea), and GlobalFoundries (U.S.).

3 COMPARISON WITH OTHER TECHNOLOGIES

FinFETs boast of several advantages as compared to other technologies in terms of static power savings, delay, scaling, output resistance and output current.[7]

	Non-FinFET	FinFET	
Active	0.1		0.1
Idle	0.9		0.9
Leakage	1	60% reduction	0.4
Dynamic	9	25% increase	11.25
Active power	0.9		1.125
Idle power	0.9		0.36
Total	1.8		1.485
Active power %	50%		76%
Idle power %	50%		24%



4 FINFET DESIGN VERIFICATION

4.1 Routing

Circuit performance is more dominated by interconnected Rs and Cs and there is significant crosstalk impact on it. Layer stacks become very heterogeneous and RC varies as much as 50 times between layers. It is also seen that significant timing variation due to layer assignment is present.Hence,interconnect optimization is becoming the centerstage of physical design.

4.2 Parasitic extraction

A trigateFinFET cross-sectional view is shown in the figure a. When the transistor is turned on, the current flows in from the contact, into the epitaxy and then a small area of source/drain extension, and through the channel, finally coming out from the other side. The total resistance is divided into several parts to characterize and build an accurate resistance model. The total resistanceconsists of two parts: the channel resistance (Rch) and the parasitic resistance (RP). The parasitic resistance contains the contact resistance (RCO), the source/drain resistance in the epitaxy region (RS/D), and the extension resistance (REXT).Studies reveal that the parasitic resistance can take up over 30% of the overall resistance of a fully turn-on transistor, which imply that a significant degradation in drive current can be attributed to RP.

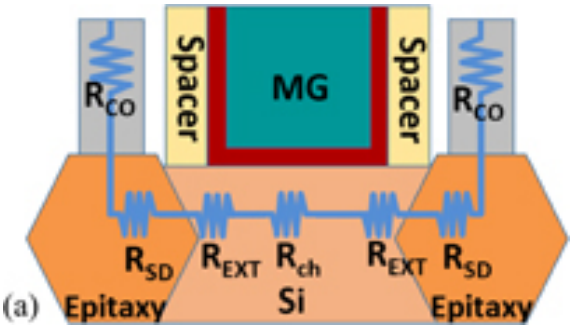


Fig 3. Resistances in FinFET

The total capacitance consists of two parts as well; the oxide capacitance and the parasitic capacitance, which consists of the contact capacitance (CCO), the epitaxial capacitance (CEpi), and the shallow trench isolation capacitance (CSTI) as shown in figure b. When channel length scales further, parasitic effects are expected to increase.

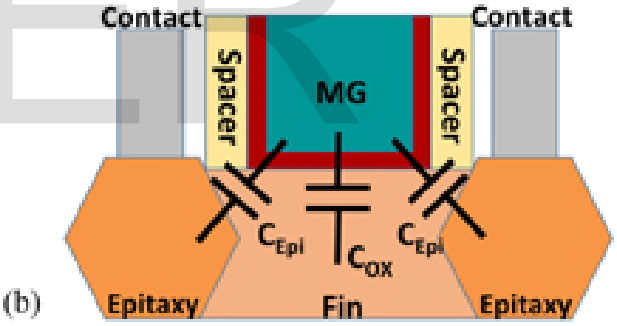


Fig 4. Capacitances in FinFET

4.3 Placement Algorithm

To reduce the current mismatch, the orientation of the sub-transistors must be properly determined. After determining the sub-transistor orientations, we want to generate a common-centroid FinFET placement while maintaining the sub-transistor orientations and maximizing the dispersion degree. When generating an m-row common-centroid FinFET placement, we first evenly distribute all sub-transistors of each transistor to the m rows such that the dispersion degree of a common-centroid FinFET placement can be effectively maximized in the subsequent steps. Given a set of ni sub-transistors of a transistor, ti, we assign ni divided by m sub-transistors into each row. If ni is less than m, we

randomly assign the sub-transistors into different rows while keeping the numbers of sub-transistors in different rows the same. Once all sub-transistors are assigned into different rows, we should consider the diffusion-sharing for transistors. We construct the diffusion graph of the sub-circuit in each row, and then we find the Euler paths on the diffusion graph.

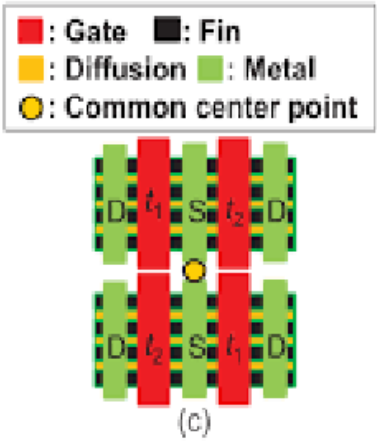


Fig 5.Common-centroid FinFET placement of a current mirror with one reference transistor, t1, and one scaled transistor, t2, where t1 and t2 are decomposed into two unit transistors, respectively.

5 BEYOND FINFETS

Though the industry will likely continue to find ways to extend CMOS finFET technology further, at some point in the not-so-distant future, making faster, lower power ICs will require more disruptive changes. For something that could be only five to seven years out, there's a daunting range of contending technologies. Improvements through the process will help, from EUV lithography and six-track design, to new materials like cobalt and ruthenium for interconnects and SiGe for channels, but new device structures will be needed before long, with researchers eyeing a smorgasbord that includes nanowires, carbon nanotubes, tunneling, monolithic 3D, spin and 2D materials.[11]

5.1 Gate all around FETs

Gate-all-around nanowires[6] could outperform finFETs at 7nm.As finFETs potential gains in speed and power start to slow at 7nm, simulations suggest gate-all-round nanowires start to look like a better option. While finFETs allowed both >40% voltage scaling (dynamic power in μW) and >20% performance gains (GHz) in moving from 14nm to 10nm, gains from moving to 7nm FinFETs will be only <30% and <15%, respectively, whereas a 7nm nanowire device should see greater >44% improvement in power and >20% improvement in performance, and similar scale gains at 5nm as well.In 7 to 10 years, IMEC researchers see the need for III-V and 2D materials, new switching mechanisms such as tunneling or spin, and optimization at the system level by various types of 3D integration, as key to continued improvements in performance.

5.2 Carbon nanotubes

Carbon nanotubes (CNTs) could potentially solve a major part of the scaling problem by their intrinsic 1nm diameter size, with IBM researchers have made progress on the key issue of making low-resistance contacts at very small scale by making atomic-level metal-to-carbon bonds at the zero-dimension point contact at the end of the CNT. CNTs also have the advantage that they can be used for both p-FETs and n-FETs, by using contact metals with different work functions. While a number of research groups grow nicely aligned CNTs on one substrate and then transfer them to the device substrate and remove the metallic CNTs from the as-grown mix, a more practical approach for uniform volume manufacturing will require depositing and patterning an already-purified, all-semiconducting CNT solution directly on the device wafer. This typically involves making a template on the substrate with conventional lithography and etch, and then functionalizing one part of that pattern to attract the CNTs from solution.

5.3 Tunnel FETs

The tendency for electrons to tunnel through a barrier at small dimensions offers a promising lower-voltage switching mechanism for scaling beyond finFETs, as it becomes harder to reduce the energy per operation of CMOS much further without reducing performance. Tunnel FETs outperform CMOS at low voltage. One option for reducing the operating voltage of future logic circuits is complementary p-type and n-type heterojunction tunnel FETs formed with a common metamorphic buffer technology. Manufacturing TFETs also will require advances in epitaxy to build the complex III-V nanocolumns, and figuring out how to etch the stacks of materials that all etch at different rat

6 CHALLENGES

FinFET is a significantly more complex device to model. Accurate FinFET parasitic extraction is more complicated. Generating good, yet compact SPICE

models is also more challenging than for planar devices.[14] The finite granularity of the fin width "W" and the limited range of freedom in channel length for a given architecture make optimizing analog as well as digital design more complex. Granted that many fins can be "ganged" together to generate a desired "W", still "L" and "W" are not exactly free continuous parameters. This is because FinFETs are 3D structures, and reining in etch variability for the high-aspect ratio processes with non-uniform pitches or locally varying pitches may be a problem. Thus FinFETs have a significant numbers of restricted design rules (RDR).

In addition to FinFET-specific challenges, the sub 14nm process node has challenges that would appear regardless of transistor technology. These include:

- The need for double patterning (using extra masks) to get features to print correctly at 20nm and below
- Layout-dependent effects, which emerge at 28nm or above and become more problematic with each new process node
- Potential 50X or more differences in resistivity between top and bottom metal layers
- Electromigration increases with each lower process node
- Complexity – how are we going to design and verify billions of transistors while meeting time to market demands?

7 CONCLUSION

The semiconductor industry is witnessing a revolutionary method of efficient design thanks to the FinFET technology. Power saving and scalability finally go hand in hand. Design and lithography procedures of FinFET are still not the most accurate; however they provide better user performance as compared to other planar transistor technologies. With this industry expected only to grow in the coming years, further research indicates that shortcomings of the FinFET can be overcome using sub 5nm nanowire technologies.

REFERENCES

[1] Moore Gordon E., "Cramming more components onto integrated circuits", Electronics, Volume 38, Number 8, April 19, 1965

[2] Hu Chenming, "FinFET and other New Transistor Technologies", seminar in University Of California, July 2011.

[3] Ming-Hung Han, "Device and Circuit Performance Estimation of Junctionless Bulk FinFETs", Electron Devices, IEEE Transactions on (Volume:60 , Issue:6), pp: 1807 – 1813, 2013

[4] "Design of High-performance digital logic circuits based on FinFET technology", International Journal of Computer Applications, pp 40-44 March 2012.

[5] Bibin Lawrence R, JencyRubia, "Review of Fin FET Technology and Circuit Design Challenges" Int. Journal of Engineering Research and Applications ISSN: 2248-9622, Vol. 5, Issue 12, (Part - 1) December 2015, pp.77-80

[6] SainiGaurav, Ashwani K Rana "Physical Scaling Limits of FinFET Structure: A Simulation Study", International Journal of VLSI design & Communication Systems (VLSICS), Vol.2, No.1, March 2011.

[7] Sameer Dwivedi, Dr.Neelam Rup Prakash," A Study on Recent Advancements in VLSI Technology using FinFETs",IJIRSET, Vol. 4, Issue 12, December 2015

[8]http://www.radioelectronics.com/info/data/semicond/fet-field-effect-transistor/finfet-technology-basics.php

[9]http://www.newelectronics.co.uk/electronics-technology/what-makes-finfets-so-compelling/56795/

[10] https://www.engineersgarage.com/articles/all-about-finfet

[11]http://semiengineering.com/pathfinding-beyond-finfets/

[12]http://www.3dic.net/a/3DICjishujieshao/2015/1025/48.html

[13]"Microelectronics Circuits", Adel Sedra and K.C Smith, 5th Edition, Oxford University Press, International Version,2009

[14] https://www.synopsys.com/designware-ip/technical-bulletin/finfet-design.html